Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Scienze Biotecnologiche e Farmaceutiche

Ciclo XXXI

Settore Concorsuale: 03/D1

Settore Scientifico Disciplinare: CHIM/08

Structural and biophysical characterization of novel GSK-3β inhibitors

Presentata da: Valentina Piretti

Coordinatore Dottorato:

Prof. Santi Mario Spampinato

Supervisore:

Prof. Andrea Cavalli

Co-supervisori:

Dott.ssa Stefania Girotto

Dott.ssa Paola Storici

Esame finale anno 2019

Al mio papà

Index

Abstract **7**

1. Introduction **8**

1.1. GSK-3β: structure, function and regulation 8

1.2. GSK-3β as a drug target for neurodegenerative diseases 13

1.3. GSK-3β as a drug target for cancer 15

1.4. Development of ATP-competitive and allosteric GSK-3β inhibitors 16

1.5. Binding kinetics in drug discovery 22

2. Aim of the research **28**

3. Materials and methods **31**

3.1. GSK-3β expression in HEK293T cells 31

3.1.1. DH5α cells transformation 31

3.1.2. Plasmid DNA purification 32

3.1.3. HEK293T cells transfection 32

3.1.4. GSK-3β purification 32

3.2. GSK-3β expression in insect cells 33

3.2.1. Ligation Independent Cloning (LIC) 33

3.2.2. Bacmid generation and purification 35

3.2.3. FuGENE® Sf9 cells transfection 36

3.2.4. Transfection optimization (PEI-MAX Sf9 cells transfection) 36

3.2.5. GSK-3β purification 37

3.3. Protein characterization 38

3.3.1. SDS-PAGE 38

3.3.2. Western Blot 38

3.3.3. Thermal Shift Assay (TSA) 39

3.3.4. Liquid chromatography-mass spectrometry (LC-MS/MS) 39

3.3.5. Fluorescent gel shift electrophoresis activity assay 40
3.4. Biophysical methods 40
3.4.1. Agarose gel electrophoresis 40
3.4.2. Protein and DNA quantification 41
3.4.3. Surface Plasmon Resonance (SPR) 41
3.4.4. Microscale Thermophoresis (MST) 43
3.4.5. Nuclear Magnetic Resonance (NMR) 43
3.4.6. Crystallization 45
3.5. Computational methods 46
3.5.1. Adiabatic bias molecular dynamics simulations 46
3.5.2. Pocketron analysis 47
3.5.3. Virtual Screening 47
4. Results and discussion 49
4.1. GSK-3β expression in HEK293T cells 49
4.2. GSK-3β expression in insect cells 52
4.2.1. Bacmid generation 52
4.2.2. Small scale expression 54
4.2.3. GSK-3β purification 55
4.3. GSK-3β characterization 59
4.3.1. Characterization of GSK-3β activity 59
4.3.2. Characterization of GSK-3β phosphorylation pattern 63
4.3.3. Characterization of GSK-3β stability 65
4.4. Validation of a novel Electrostatics Adiabatic Bias Molecular Dynamics (elABMD) method to predict protein-ligand unbinding kinetics 67
4.4.1. Determination of k_{off} values with SPR 69
4.4.2. TSA on ATP-competitive inhibitors 74

4.4.3. Crystallization of ATP-competitive inhibitors C22, C44 and C50 76

4.4.4. Computational prioritization and comparison with experimental SPR data 84

4.5. Characterization of GSK-3β allosteric inhibitors 20g and 20f 88

4.5.1. Thermal Shift Assay 89

4.5.2. Microscale Thermophoresis 89

4.5.3. Crystallization trials 90

4.6. Identification of novel GSK-3β allosteric inhibitors 93

4.6.1. Identification of allosteric pockets 93

4.6.2. Virtual screening 96

4.6.3. MST on Virtual Screening hits 97

4.6.4. Fragment-based NMR screening 99

5. Conclusions and perspectives 103

References 107

Abstract

The deregulation of GSK-3β (glycogen synthase kinase-3β) is involved in the pathogenesis of diverse diseases, such as cancer, diabetes and neurodegenerative disorders. Therefore, GSK-3β has become an attractive target for the design and development of new inhibitors for pathologies that present many limitations in therapeutic treatment. In the present thesis, we report a fast and efficient protocol for the expression of GSK-3β in insect cells with the baculovirus system, and employ purified GSK-3β in two different drug discovery projects. The first project describes the perspective validation of a novel computational method based on adiabatic bias molecular dynamics (ABMD) that aims to simulate protein-ligand unbinding events. Surface plasmon resonance (SPR) experiments performed on a series of GSK-3β inhibitors confirmed a coherence between predicted and experimental k_{off}, highlighting the potential of this method for the calculation of residence times in hit-to-lead and lead optimization phases of drug discovery programs. Three novel X-ray crystal structures of pyrazine inhibitors in complex with GSK-3β are reported, providing further details in the binding mode of ATP-competitive GSK-3β inhibitors. The second project involves the structural and biophysical characterization of known GSK-3β allosteric inhibitors, and, in parallel, the discovery of novel allosteric modulators through computational and biophysical techniques. Herein, we describe the identification of promising molecules that have been selected by virtual screening and microscale thermophoresis (MST) analysis that display affinity to GSK-3β in the micromolar range. Taken together, our results provide useful insights for future rational drug design and discovery of small selective GSK-3β inhibitors.

1. Introduction

1.1. GSK-3β: structure, function and regulation

GSK-3 (glycogen synthase kinase-3) was first discovered in 1980 as one of the protein kinases able to phosphorylate and inhibit glycogen synthase, the enzyme involved in the last step of glycogen synthesis (1). Under basal conditions, glycogen synthase is maintained in its inactive form due to GSK-3 phosphorylation. In presence of high glucose and insulin availability, GSK-3 activity is inhibited via the action of insulin through PI3/PKB pathways, resulting in the activation of glycogen synthase and in the formation of glycogen (2). Besides its role in glycogen metabolism, GSK-3 is involved in many different physiological pathways such as cell cycle, development, neuroprotection and apoptosis (3). It phosphorylates and regulates an array of proteins involved in transcription (c-Jun, NFAT, CREB), cell cycle (cyclin D1), translation initiation (eIF2B) and Wnt pathway (axin and β-catenin) (Figure 1. 1) (4). Many of these proteins are subsequently degraded as a direct result of this phosphorylation.

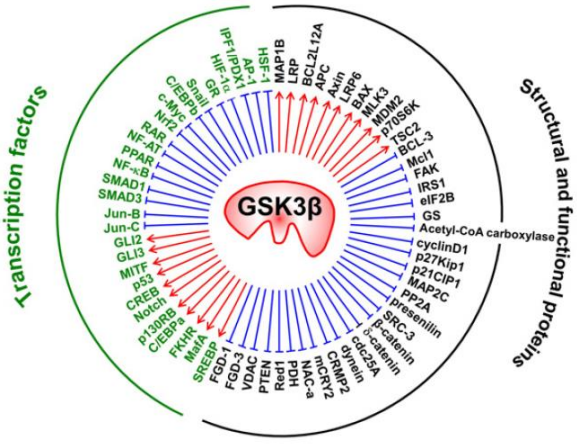


Figure 1. 1. Substrates of GSK-3β and proteins that interact with it (5)

GSK-3 exists in two isoforms named GSK-3α (51 kDa) and GSK-3β (47 kDa) which are expressed in most tissues and are encoded by genes GSK3A and GSK3B, respectively (6). The two isoforms share 98% similarity and have nearly identical catalytic domains, but GSK-3α contains a large glycine-rich N-terminal region that is not present in GSK-3β and that is responsible for GSK-3α cytoplasm localization. They also diverge in the C-terminal sequence, showing only 36% identity in the last 76 residues (Figure 1. 2). An alternative splicing variant of GSK-3β named GSK-3β2 has also been reported, containing 13 amino acids inserted in the catalytic domain. GSK-3β2 isoform is enriched in neurons and there is preliminary evidence that this variant exhibits differential substrate specificity, although its role remains unclear (7).

Although the two isoforms share many substrates, they are not functionally identical and show some differences in functional effects, substrate specificity and expression regulation. For many years they have been considered as identical twins since all known GSK-3 inhibitors inhibit both isoforms, making difficult to identify differential effects (5). GSK-3α is less expressed in the reticulocytes, nerves, ovary and skin, while it is expressed at higher levels in the appendix, whole blood and pituitary gland. GSK-3β is less expressed in the reticulocytes, lymph nodes and pancreas and expressed at higher levels in blood NK cells and bone marrow granulocytes (8). GSK-3β is mainly localized in the cytosol, but lower amounts are expressed in the nucleus and mitochondria, and its regulatory function in the mitochondrial cell death pathway has been elicited by a variety of stress conditions shown in neuronal cells (9). GSK-3β is the predominant isoform in the brain (10). The advent of knock-out and knock-down methods has allowed the identification of different actions of GSK-3α and GSK-3β: it was observed that GSK-3β knock-out mice are embryonically lethal, while GSK-3α knock-out mice are viable and relatively normal, demonstrating that the two isoforms do not have the same physiological effect. Moreover, there is a growing evidence that some substrates are differentially phosphorylated by the two isoforms, such as the p62 nucleoporin protein and the transcription factors early growth response-1 and Smad3/4 (11).

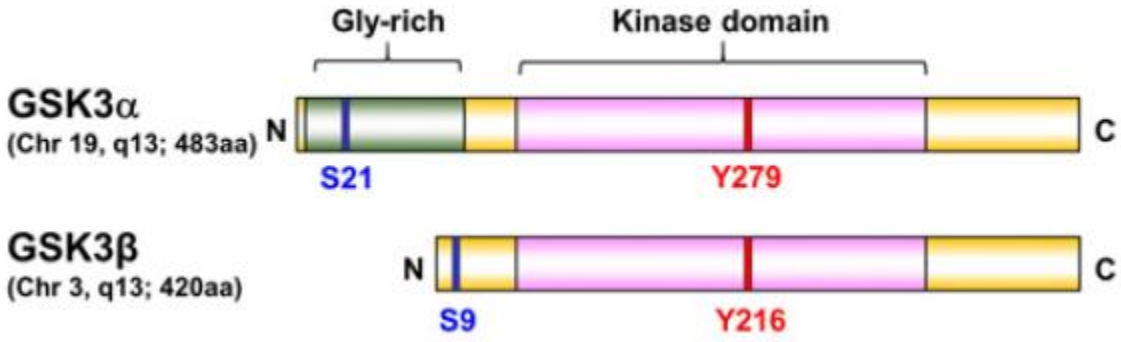


Figure 1. 2. Comparison of the structural and functional domains of the two GSK-3 isoforms (5)

The majority of the studies have been mainly focused on GSK-3β in the field of oncology and neurodegenerative disorders. This is probably due to the involvement of GSK-3β in the Wnt/β-catenin pathway, which is responsible for generating the most prevalent oncogenic signaling (5). For this reason, this work will entirely focus on the GSK-3β isoform.

GSK-3β is formed by the typical two-domain kinase fold: an N-terminal β-strand domain (residues 25-138) consisting of seven antiparallel β-strands, and a C-terminal α-helical domain (residues 136-343). The ATP binding site is located at the interface between the two domains and is adjoined by a glycine-rich loop and hinge region. The activation loop (residues 200-226) is well ordered and runs along the surface of the substrate binding groove, and the C-terminal residues (344-382) are located outside the core kinase fold (12).

GSK-3β differs from most kinases for three main reasons: it is constitutively active, its substrates need to be already phosphorylated by another kinase, and it is inhibited, rather than activated, in response to stimulation of the insulin and Wnt pathways (11). The substrate specificity is unusual in that efficient phosphorylation of most of its substrates requires the presence of another phosphorylated residue (termed the *primed phosphate*) located four aminoacids C-terminal to the site of GSK-3β phosphorylation. The most common target for phosphorylation is the sequence S/T-X-X-X-S/T(P), in which GSK-3β phosphorylates a Ser/Thr four residues N-terminal to a pre-phosphorylated Ser/Thr. The binding site for the priming phosphate has been recently identified, and it was found to contain three crucial basic residues: Arg96, Arg180 and Lys205 (1).

GSK-3β has two phosphorylation sites that regulate its catalytic activity: phosphorylation of Ser9 inactivates the protein, whereas phosphorylation of Tyr216 increases its catalytic activity. Phosphorylation of Ser9 is the major regulatory checkpoint in most pathways (13) and results in the N-terminal domain of GSK-3β interacting with its phosphate binding pocket, preventing recognition of primed substrates (10). Ser9 is phosphorylated by many protein kinases, including protein kinase A (PKA), integrin-linked protein kinase (ILK) and Akt, and by many physiological processes, such as Insulin/IGF1 and NGF (14). Thus, in response to insulin, the inhibition of GSK-3β promotes the dephosphorylation and activation of glycogen synthase, contributing to the stimulation of glycogen synthesis. GSK-3β can also be phosphorylated at Ser9 by RSK, a protein part of the mitogen-activated protein kinase (MAPK) cascade, providing a route for the inhibition by growth factors and other signals that activate this pathway. Ser9 can be phosphorylated by p70 ribosomal S6 kinase-1 (S6K1) and by cyclic-AMP-dependent protein kinase PKA in response to agonist that elevate the intracellular concentration of cAMP (15).

The structural basis of the inhibition of GSK-3β by phosphorylation of Ser9 were first investigated by Dajani et al. in 2001 (16). GSK-3β activity was analyzed in the presence of a peptide based on the N-terminal segment in which Ser9 is located (residues 3-12). A significant dose-dependent inhibition of the protein activity was observed with the phosphorylated peptide, while no inhibition occurred with the identical unphosphorylated peptide. These results suggested an autoinhibitory effect in which the phosphorylated N-terminal segment binds in the substrate-binding cleft as a pseudo-primed substrate, hindering the binding of other substrates and thus diminishing primed-substrate phosphorylation by GSK-3β. However, phosphorylation of Ser9 does not cause absolute inhibition of the activity: the Ser9 tail appears to be in a dynamic state of phosphorylation and dephosphorylation. At high concentrations, primed substrates are able to displace pSer9 and be phosphorylated by GSK-3β once again (Figure 1. 3) (11).

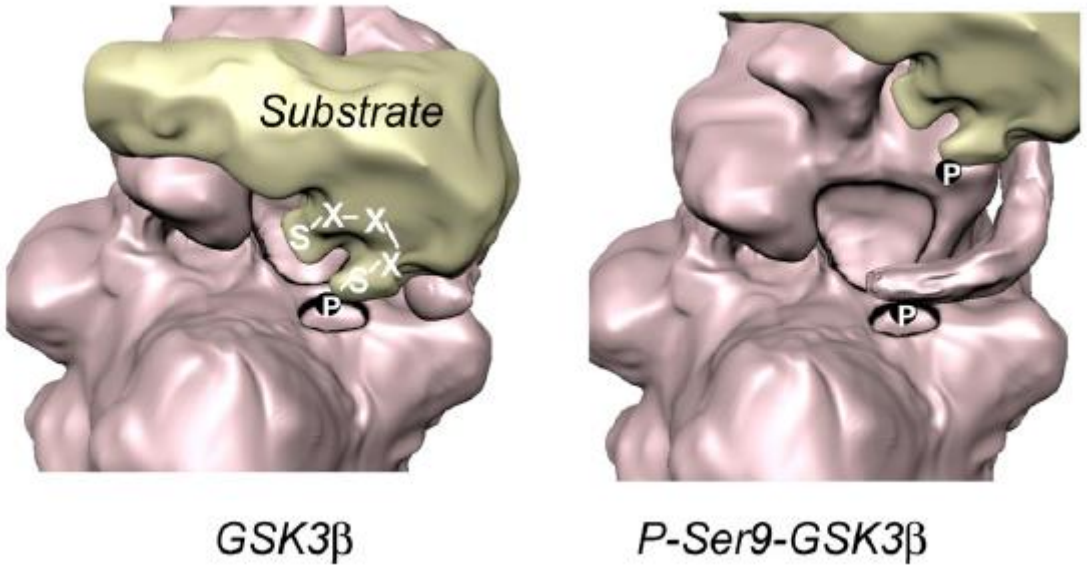


Figure 1. 3. Representation of how Ser9 phopshorylation inhibits GSK-3β phosphorylation of primed substrates. On the left, a primed substrate is bound to the primed substrate binding domain of GSK-3β. On the right, pSer9 in the N-terminal tail of GSK-3β inhibits the association of primed substrates (11).

Besides Ser9, GSK-3β is also phosphorylated on Tyr216, but its regulatory function remains unclear. Phosphorylation of this residue is crucial for proper folding of the catalytic domain, and it occurs through autophosphorylation during synthesis of the GSK-3β polypeptide. As such Tyr216 is likely to be constitutively phosphorylated to high stoichiometry (10).

Crystal structures containing unphosphorylated Tyr216 (16) (12) show that while the tyrosine side chain points towards the N-terminal lobe, the tyrosine ring is packed against Val214. The presence of pSer9 precludes this positioning and instead it points outward and adopts an identical conformation of the unphosphorylated form. The increase in GSK-3β activity when Tyr216 is phosphorylated is explained by the fact that in order to accommodate the binding of the substrate, Tyr216 must move from its unphosphorylated position and break its interaction with Val214, therefore requiring extra energy for the rearrangement (13) (12). Crystallographic studies suggest that Tyr216 activates the protein through interaction with Arg220 and Arg223, thus stabilizing the activation loop in an active conformation and allowing complete substrate accessibility (17) (18).

1.2. GSK-3β as a drug target for neurodegenerative diseases

GSK-3β is a point of convergence for multiple signaling pathways and therefore plays a crucial role in regulating the pathogenesis of diverse diseases. A growing number of studies have demonstrated that the deregulation of GSK-3β activity contributes to the pathogenesis of diabetes, cancer, neurodegenerative disorders, chronic inflammatory and immunological diseases (5). GSK-3β has become one of the most attractive targets for medicinal chemists to discover, design and synthetize new selective and potent inhibitors to treat diseases that currently have many limitations in therapeutic treatments – such as Alzheimer's disease, Parkinson's disease and Amyotrophic Lateral Sclerosis. All of these pathologies are characterized by the loss of neurons in particular regions of the nervous system and unknown etiology. From a molecular point of view, these neurodegenerative diseases share, in addition to neuronal death, deficits in neurotransmitter systems and protein misfolding with aberrant aggregation.

Alzheimer's disease (AD) is a neurodegenerative disease that leads to neurodegeneration and severe cognitive impairments. The neuropathology of AD is characterized by extracellular plaques of aggregated amyloid-β peptide (Aβ) and intracellular neurofibrillary tangles containing hyperphosphorylated tau (τ). The formation of neurofibrillary tangles leads to neurodegeneration and neuronal deaths (7). The evidence that GSK-3 plays a pivotal role in AD and that its deregulation accounts for many of the pathological hallmarks of the disease, has led to the development of the "GSK-3 hypothesis of AD". In 1992 it was discovered that GSK-3 contributes to the abnormal phosphorylation of tau protein (11), a neuronal protein involved in microtubule stabilization and axonal transport. The phosphorylation of tau takes place in the regions surrounding the microtubule binding domain, preventing the interaction of tau protein with microtubules and therefore affecting microtubule stabilization and dynamics (19). GSK-3 also promotes the aggregation of amyloid-β peptide, which leads to amyloid plaques (20). The exposure of neurons to Aβ increases GSK-3β activity through the inhibition of PI3-kinase signaling (21).

The finding that GSK-3 plays a central role in the pathophysiology of AD has led many researchers to study the effect of GSK-3 on cognition in rodent models and in patients with AD. It has been observed that the suppression of overexpressed GSK-3β reduced tau phosphorylation, reactive gliosis and neuronal death. The GSK-3-tau interaction was further investigated in Tau-P031L mice with increased GSK-3β expression. These mice displayed impaired novel object recognition and memory prior to the deposition of tau aggregates, suggesting that early tau pathology due to increased GSK-3β activity may be the cause of synaptic deficits underlying the cognitive impairments (21). Moreover, modulation of GSK-3β signaling pathway showed neuroprotective effects in transgenic mice by regulating amyloid precursor protein (APP) maturation and processing, further supporting the notion that GSK-3β may be a suitable target for the treatment of AD (22).

Recently, experimental evidences have shown that alterations of GSK-3β are present on patients with Amyotrophic Lateral Sclerosis (ALS) (23). ALS is a progressive neurodegenerative disease that affects neuronal cells in the brain and the spinal cord, leading to the degeneration and death of motor neurons. Although the underlying cause of ALS is still unclear, a wide range of mechanisms is thought to be implicated in ALS pathogenesis, including mitochondrial dysfunction, protein misfolding, oxidative stress and aberrant growth factor signaling. Increased levels of protein kinases, including GSK-3α and GSK-3β, have been found in the spinal cord of patients with sporadic ALS. It has also been observed that in an *in vitro* familial ALS model, in which motor neuron cells have the G93A mutant superoxide dismutase (SOD1) gene, GSK-3β activity was increased, and its inhibition prevented motor neuron cell death (24). Inhibition of GSK-3β activity has been also evaluated in an ALS mouse model, and the effect on the onset of symptoms and disease progression was monitored by observing the molecular changes in the spinal cord of ALS mice: Koh and colleagues (24) demonstrated that inhibition of GSK-3β activity with small molecule AR-A014418 delayed the onset of symptoms, delayed the disease endpoint, suppressed disease progression and prolonged time until motor failure. Moreover, the neuroprotective effect of GSK-3β inhibition was confirmed in both morphological and behavioral tests.

These recent studies collectively suggest that the gain of function of GSK-3β could be one of the potential pathogenic mechanisms of ALS and have led to hypothesize that the inhibition of GSK-3β activity could represent a promising therapeutic strategy (23).

It has been observed that GSK-3β facilitates numerous apoptotic conditions involved in Parkinson's Disease (PD) pathogenesis, including mitochondrial dysfunction, oxidative stress, protein aggregation and the inflammatory response. PD is a neurodegenerative disorder characterized by a gradual and progressive loss of dopaminergic neurons and by intracellular inclusions known as Lewy bodies expressed in surviving neurons of the substantia nigra. Lewy bodies are mainly composed by α-synuclein, a presynaptic neuronal protein abundantly expressed in the nervous system. The involvement of GSK-3β in mitochondrial dysfunction is highlighted by the fact that it can regulate cell survival and apoptosis by controlling mitochondrial complex I activity and ROS production, and this can be reversed by GSK-3β inhibitors. GSK-3β phosphorylates α-synuclein and decreases its protein expression, preventing cell death, therefore the inhibition of GSK-3β may be neuroprotective to dopaminergic neurons by attenuating the toxicity of α-synuclein overexpression (9). 6-Hydroxydopamine (6-OHDA) is a PD mimetic widely used to model PD in vitro and in vivo. Results obtained using human dopaminergic neuronal cell line SH-SY5Y showed that GSK-3β is a critical intermediate in pro-apoptotic signaling cascades associated with PD, and that the blockage of GSK-3β prevented 6-OHDA-induced cleavage of caspase-3, DNA fragmentations and cell death (22). The cellular and molecular mechanisms of the protective effect of GSK-3β inhibition on dopaminergic neurons require further elucidation, and could provide a potential efficient target for treating PD by blocking the pathogenic pathway.

1.3. GSK-3β as a drug target for cancer

The role of GSK-3β in cancer is largely investigated but remains controversial, since its activity has been associated with both tumor progression and tumor progression inhibition. GSK-3β has shown to play a positive role in cancer cell proliferation and survival, and is overexpressed in various

tumor types including colon, liver, ovarian and pancreatic cancer (25). Whether GSK-3β functions as a tumor suppressor or a tumor promoter seems to depend on the tissue type of the tumor. In some cases, the suppression of its activity has been associated with cancer progression (8). It has been shown that GSK-3β phosphorylates and targets for proteosomal degradation many proto-oncogenic proteins, such as c-Jun, cyclin D1 and c-Myc, and negatively regulates cell cycle regulator via the Wnt/β-catenin and Hedgehog pathways. These evidences have led to the credentialing of GSK-3β as a tumor suppressor (6)(26).

In other cases, however, GSK-3β has been associated with tumor progression: it has been demonstrated that the inhibition of GSK-3β results in apoptosis induction through decrease expression of NFkB target genes Bcl-2 and XIAP in chronic lymphocytic leukemia and pancreatic cancer cells (27). GSK-3β has also been identified as a positive regulator of proliferation, survival and chemoresistance of bladder and renal cancer cells. In bladder cancer patients, GSK-3β aberrant nuclear accumulation is associated with high-grade tumors, metastasis and worst survival rate (28).

The investigation of the mechanisms that underlie opposite actions of GSK-3β in different types of cancer may provide opportunities to understand the mechanistic cellular changes and to identify the ones that could safely benefit from GSK-3β inhibitors administration.

1.4. Development of ATP-competitive and allosteric GSK-3β inhibitors

The increasing evidence that GSK-3β contributes to the pathology of several prevalent diseases has generated much interest in applying GSK-3β inhibitors therapeutically. The first GSK-3 inhibitor to be identified was the lithium cation, which remains the most widely used experimentally and clinically (21). It inhibits GSK-3β directly by competition with magnesium ions and indirectly by enhanced serine phosphorylation and autoregulation (29). Lithium has been extensively tested in different AD animal models in which it blocked amyloid precursor protein (APP) deposits, decreased tau hyperphosphorylation and Aβ levels (22). The utility of lithium and therapeutic promise of GSK-

3β inhibitors have led to the development of many selective inhibitors with great chemical diversity, most of which are ATP-competitive inhibitors.

Among the most frequently used are maleimide derivatives (30), indole derivatives, paullone derivatives and indolocarbazoles. Paullones are fused tertracyclic compounds that inhibit both GSK-3 and CDKs within the nanomolar range. Derivatives alsterpaullone and kenpaullone have shown to reduce tau phosphorylation and decrease Aβ production in cells, respectively. As shown in Figure 1. 4, alsterpaullone establishes two hydrogen bonds with Val135 and one interaction between the nitro group and the side chain of Lys85 of GSK-3β (30), and was shown to prevent neuron cell death in response to variety of insults including trophic deprivation and mitochondrial stress (29). Similar compound 1-azakenpaullone and its derivative cazpaullone have been recently characterized as selective GSK-3 inhibitors. Other potent ATP-competitive inhibitors include indole derivatives such as indirubins, which have long been used in traditional Chinese medicine for the treatment of leukemia. Additional compounds were recently described as dual CDK/GSK-3 inhibitors, such as purine derivatives pyrazolo [3,4-*b*] quinoxalines (31), the pyrazolo [3,4-*b*] pyridine ring system (32) and the the9-oxo-thiazolo [5,4-f] quinazoline-2-carbonitrile derivatives (33). Aloisines are a different class of dual CDK/GSK-3 that inhibit both kinases in the sub-micromolar range: aloisine A, the most potent of the analogs testes, showed anti-proliferative effects in differentiated postmitotic neurons (29). The amino thiazole AR-A014418 was identified through high-throughput screening and displayed GSK-3 specificity in a panel of 26 kinases, as well as a neuroprotective effect in apoptotic conditions (34). A novel series of pyrazine analogues was described by Berg and colleagues and characterized as potent ATP-competitive inhibitors showing good solubility and permeability in a blood brain barrier assay (35).

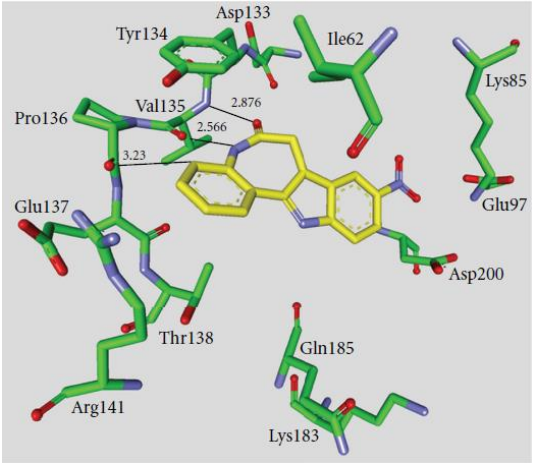


Figure 1. 4. Alsterpaullone in the ATP-binding pocket of GSK-3β. PDB code 1Q3W. (30)

Despite wide research for the past 20 years, most GSK-3β ATP-competitive binders display poor to moderate selectivity when tested against the larger human kinome at a concentration comparable to cellular exposure, or physicochemical properties unsuitable for use in *in vitro* systems or *in vivo* models (4). Kinase selectivity is, in fact, one of the main challenges in the design of protein kinase inhibitors due to the high degree of identity of the ATP binding pocket (36), and because of that ATP-competitive inhibitors could offer severe side effects in chronic treatment.

Structural analysis of the ATP competitive inhibitors scaffolds can help the development of more selective inhibitors: it has been observed that hydrogen bonds with the backbone atoms of Asp133 and Val135 are a key interaction to enhance affinity to GSK-3β that, however, does not provide selectivity over other kinases. In many complexes, Pro136 appears to strengthen the interaction with the protein backbone. The interaction of an inhibitor with the region characterized by aminoacids Lys85, Glu97 and Asp200 holds potential to increase activity and selectivity for GSK-3β, and can be mediated by water molecules or be established by direct contact (30).

The development of selective ATP-competitive GSK-3β inhibitors is a dynamic research field, but alternative approaches in the design of novel inhibitors are increasingly being pursued, in the search for safer and more selective drugs. This challenge may be achieved considering allosteric ligands, which bind in a site distinct from the ATP pocket. Advantages of allosteric inhibitors include minimal off-target pharmacology and the potential to overcome mutation-associated drug-resistance.

Moreover, this type of inhibitors may not need to show nanomolar affinity in order to compete with the high intracellular ATP concentrations, allowing the identification of weak binding inhibitors. However, due to the hydrophobic properties of most allosteric pockets, most allosteric inhibitors are lipophilic compounds with poor bioavailability and solubility (37).

The field of allosteric kinase inhibitors has grown in the past few years since the FDA-approval of trametinib as the first small-molecule kinase allosteric inhibitor (37) which selectively inhibits mitogen-activated protein kinase 1 and 2 (MEK1 and MEK2). Despite the therapeutic potential, during the past years only a few different families of non-ATP competitive inhibitors have been reported: in 2002 Martinez et al. (38) reported some thiadiazolidinediones (TDZDs) as the first non-ATP competitive GSK-3β inhibitors. It was hypothesized that these compounds recognize and interact with the oxyanion binding site. Among them, tideglusib (NP-12) was designed for the treatment of AD and progressive supranuclear palsy, and so far, it is the only non-ATP competitive GSK-3β inhibitor that has reached clinical trials. Since the discovery of TDZDs, extensive chemical SAR studies on the thiadiazolidindione scaffold were performed by varying the heteroatoms present in the central ring and the substituents. Following this approach, the class of 5-Imino-1,2,4-Thiadiazoles (ITDZs) was described as the first substrate-competitive GSK-3β inhibitors (39). These compounds bind reversibly to GSK-3β through (I) hydrogen bonding between the backbone oxygen of Phe67 and the imino charged group, and (II) aromatic S-π interaction between the aromatic ring of Phe67 and the thiadiazole nucleus (Figure 1. 5). ITDZs showed to be cell permeable and able to penetrate the CNS in an in vitro model.

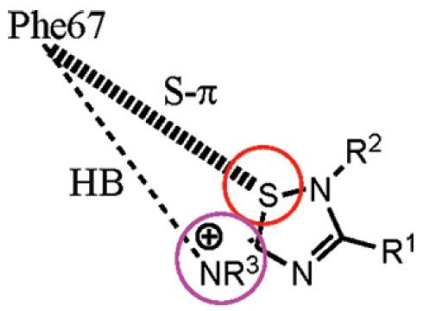
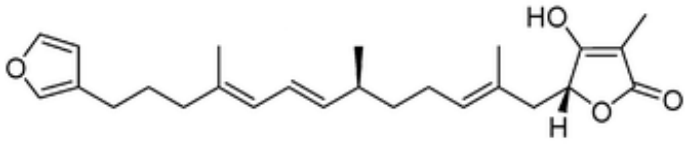


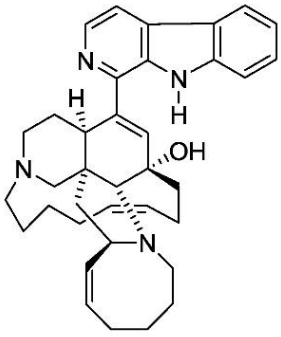
Figure 1. 5. Key structural interactions for the binding of ITDZs to GSK-3β (39)

Another family of non-ATP competitive inhibitors is halomethylketones (HMKs), which have been classified as the first irreversible GSK-3β inhibitors (40). The mechanism of action involves the formation of an irreversible covalent sulfur-carbon bond between the HMK moiety and the residue Cys199.

Recently, a number of GSK-3β inhibitors derived from marine organisms has been identified, including meridianines, indirubines, isoflavones and manzamines, highlighting the potential of marine products in the development of effective inhibitors. Among these, the β-carboline alkaloid manzamine A (41) was isolated from the Indonesian sponge *Acanthostrongylophora*. Palinurin (42) is a linear furanosesquiterpene composed by a hydrophilic, negatively charged tetronic ring, and a hydrophobic sesquiterpene fragment. It was extracted from the Mediterranean sponge *Ircinia variabilis*. Both compounds were reported to efficiently decrease tau phosphorylation as cell permeable non-ATP competitive inhibitors.



Palinurin

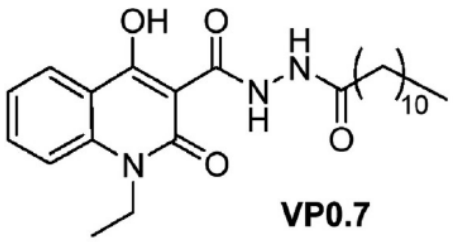


Manzamine A

Experimental data and MD simulations suggested that palinurin acts through the binding to an allosteric site located at the N-terminus of the protein, discovered as pocket 5 in a previous study (36). Molecular modeling studies have also revealed a potential allosteric site on GSK-3β for manzamine A (43). The results showed that manzamine A fits well into the ATP-noncompetitive pocket located in proximity of the activation pocket formed by Arg96, Arg180 and Lys205, establishing three polar interactions and several nonpolar interactions. Despite the fact that manzamine A and palinurin

constitute a promising scaffold for the development of potent and selective GSK-3β inhibitors, further studies were limited by their chemical structure.

In a recent paper, Palomo and colleagues have investigated the druggable binding sites on GSK-3β able to be modulated by small-molecules using the geometry-based algorithms fpocket and hpocket (36). They detected seven conserved cavities: three of them were well-known binding sites (ATP-, substrate- and axin/fratide binding site). The other four additional binding sites were identified to be non-ATP sites. They also assayed their in-house chemical library and found that compound VP0.7 acts as a noncompetitive inhibitor of ATP binding, as demonstrated by the lack of sensitivity to increased ATP concentration on enzymatic inhibition. The ATP-independent interaction was further supported by a docking study of the quinolone moiety.



VP0.7

In 2012 Zhang and colleagues (44) discovered a novel scaffold of benzothiazinones (BTOs) through virtual screening approach. Over fifty BTOs were synthetized with different alkyl, aryl and alkylaryl as R group and side chain (45). Among them, two structurally similar novel compounds (BTO-5h and BTO-5s) which acted as allosteric modulator and substrate competitive inhibitor, respectively, were identified. However, the hydrazine contained in BTO-5h was classified as genotoxic and was recommended as structural alert. Aiming at replacing the potential genotoxic structure, a second series of novel BTOs derivatives was synthetized: the most promising compound, 20g, acted as non-ATP competitive GSK-3β inhibitor and showed suppression efficacy on ovarian cancer (46).

Currently, information regarding the side effects of GSK-3β inhibitors is limited since clinical trials have evaluated only a small number of compounds, and lithium is the only currently approved inhibitor for clinical use. The high number of substrates that are phosphorylated by GSK-3β may limit

its feasibility as a therapeutic target because of the potential disruption of many cellular processes. Many concerns have been raised regarding the potential toxicity of GSK-3β ranging from hypoglycemia to tumorigenesis and neuron deregulation (29). Moreover, these inhibitors are of distinct chemical structures and thus differ in their bioclinical and pharmacological properties such as absorption, distribution and metabolism. Identification of the range of GSK-3β inhibition that is both efficient for treatment of disease and results in minimal detrimental effects to normal tissues and organs will be fundamental for the clinical application of these inhibitors (5) (29). For this reason, a partial inhibition of GSK-3β may be optimal for reducing its self-activating mechanisms in disease processes, while allowing it to fulfill its many other cellular actions (11). A mild inhibition could provide sufficient insulin sensitization without elevation of β-catenin levels and the minimization of significant mechanism-based toxicities (30). In this scenario, allosteric GSK-3β inhibitors would provide important benefits in therapeutic use, including better kinase selectivity and a lower value of IC50.

1.5. Binding kinetics in drug discovery

A drug is effective as long as it remains bound to its target. Consequently, most early-stage drug discovery projects have focused on optimizing the affinity and selectivity of the drug-receptor complex, considering affinity as an appropriate surrogacy for *in vivo* efficacy (47). Although many effective drugs have been discovered on this basis, recent studies show that the kinetics of drug-receptor binding could be, in some cases, more important than affinity in determining drug efficacy (48). In a single step drug binding model, the ligand (L) reversibly binds to its receptor (R), forming a complex (RL) (Figure 1. 6).

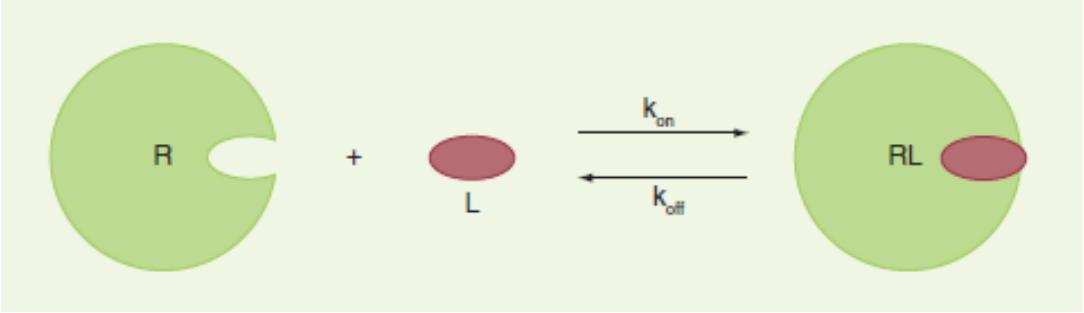


Figure 1. 6. Static lock-and-key model of receptor-ligand binding (49).

The rates at which this complex forms and dissolves are the so-called association (k_{on}) and dissociation rate (k_{off}). Molar concentrations of the drug, the receptor, and complex, do not change once the system is in equilibrium. The affinity of this interaction is measured by the equilibrium dissociation constant K_D, which provides the extent to which the drug is bound in equilibrium (47) (Equation 1).

$$KD = \frac{[R]eq\ [L]eq}{[RL]eq}$$

Equation 1

Thermodynamics and kinetics of binding are related by Equation 2 in equilibrium.

$$KD = \frac{Koff}{Kon}$$

Equation 2

In *in vivo* systems, drugs and targets usually encounter each other in a compartment (e.g. the cell, or the cell surface) in which the compound diffuses in and out, and it is actively extruded by drug pumps. Therefore, while the concentration of the target remains constant, the concentration of compounds is continuously changing. In this system, there has been increasing recognition that *in vivo* drug target interactions are not defined by equilibrium conditions and that K_D may not be an adequate descriptor of *in vivo* efficacy (50).

The term drug-target residence time (RT) has been coined by Copeland in 2006 to describe the temporal duration of the drug-target complex under different conditions, as "*the experimental*

measure of the lifetime of a drug-target complex, reflected by its dissociation rate (k_{off})" (51). The definition of residence time is independent of the environment of the study, so that the receptor-ligand complex will have the same mean lifetime in the equilibrium and non-equilibrium scenario. Residence time (τ) is expressed as the reciprocal of the dissociation rate constant ($\tau = 1/k_{off}$) or the dissociative half-life ($t_{1/2} = 0.693/k_{off}$).

Although the association rate can have an important impact on pharmacology, attention has been mainly focused on the dissociative process (and therefore the residence time). The association step is mainly gated by diffusion-limited collisional processes, or by the physicochemical steps such as absorption and distribution, thus is generally difficult to manipulate through medicinal chemistry. On the other hand, dissociation rate is entirely determined by the drug-target interactions and *in vivo* it determines the duration of target occupancy (52). Although there are several factors that drive drug residence time which are often hard to characterize, it is considered a useful parameter for drug optimization. It is thought to represent a marker of drug clinical efficacy, following the idea that the longer the drug occupies the target, the more profound the drug may exert its effects. Slow dissociation of protein-drug complexes is expected to require lower drug concentrations, thus minimizing off-target toxicity effect and increasing therapeutic index (53).

While the concepts that underlie rational optimization of binding affinity are relatively well understood, the same is not true for binding kinetics, for which less is known about the molecular determinants. Drug residence time has been recognized as an important parameter for the development of effective inhibitors, but the lack of understanding of structure-kinetics relationships (SKR) precludes the rational optimization of this property and the assessment of how kinetic signatures influence the downstream process in drug discovery. Many research groups have tried to apply SKR analogous to the classical structure-activity relationships (SAR) used by medicinal chemists to investigate the structural basis of the observed kinetic profile. Examples of such structure-kinetic analyses include HIV-1 protease inhibitors (54) and compstatin analogs (55). These studies, however, did not provide information about the structural basis of slow off-rates that could further be applied in prospective drug design. One approach to increase drug residence time is to increase the drug size and rigidity, or to change interactions between the drug and the fluctuating parts of the

receptor that often play a crucial role in drug binding. Harnessing electrostatic interactions requires a subtle approach, since modification of the charge of a drug can cause significant and deleterious changes in ADME properties; this can be achieved by either modulating amine basicity and carboxylates and N-*aryl*-sulfonamides acidity, or by altering the distribution of charges in the drug, thereby modulating particular interactions (48). The dissociation rate is defined by the difference between the ground state of the active enzyme-inhibitor complex and the transition state for the exit of the inhibitor. Modification of kinetics can therefore be performed in two ways: one way is to stabilize or destabilize the transition state while keeping the bound and unbound species invariant, the other is to stabilize bound state therefore increasing the affinity and decreasing k_{off}. (56).

As there is a growing interest in drug residence time, there is also growing interest in assays to reliably measure it. Experimental techniques for measuring binding kinetics can be divided into methods using a label for detection of ligand binding (such as radioactivity or fluorescence assays), label-free assays (such as biosensor technologies) and enzyme activity assays. Radioactive labeling of compounds implies the replacement of specific atoms by their radioactive isotope. Kinetics with radioligand binding can be measured with a direct method or indirectly by competition experiments. In the direct method, the receptor is pre-incubated with a known concentration of the radioligand, then bound and unbound radioactive molecules are separated either by filtration or by washout (Georgi 2017). K_{on} is calculated by performing association experiments at different radioligand concentrations. Alternatively, kinetics of unlabeled drugs can be measured by competition displacement by a radioligand of known affinity (47). These assays are often employed for kinetic binding measurements with G-protein coupled receptors (GPCRs), as they are compatible with readily available samples such as cells, tissue slices and plasma membrane preparations. They allow the direct measurement of rate constants k_{on} and k_{off}, but the individual radiolabeling of ligands is expensive, labor intensive, and provide low throughput. Fluorescent labels are an alternative to the use of radioligands and include time-resolved fluorescence resonance energy transfer (TR-FRET), fluorescence anisotropy and intrinsic fluorescence. They provide advantages in terms of sensitivity, safety and multiplexing, but show some drawbacks that can result from background fluorescence, photobleaching and steric hindrance by the bulky label. The most widely used label-free technique is

surface plasmon resonance (SPR), which is well suited for the kinetic characterization of purified soluble proteins. It offers the advantage of a direct observation of binding kinetics with low material consumption, without the presence of competitive molecules, such as ATP or substrates. Enzymatic activity assays are an alternative label-free method to determine the binding kinetics from enzyme activity. They include, for example, jump dilution assay, in which the dissociating kinetics is calculated by pre-incubating with high ligand concentration and diluting a hundredfold afterwards.

Recently, several groups have used modeling and structural biology to understand the main drivers of slow off-rates. Molecular dynamics (MD) studies on full binding and unbinding pathways have recently emerged, allowing the characterization of structural and dynamic determinants in atomic detail of ligand-receptor coupling with femtosecond temporal resolution. MD is capable of providing quantitative binding estimates with high accuracy, but it is typically being considered too computationally expensive.

Due to increases in computational power and the design of specialized hardware, only recently it has become possible to simulate the full process of spontaneous ligand-receptor interaction – that typically occurs on the microsecond timescale – in atomic detail, providing access to information on binding mechanism which have been difficult to assess experimentally (48). A variety of *enhanced sampling* or *advanced sampling* methods have been applied to reframe the sampling problem in novel ways and observe long-time behaviors.

The three main families of enhanced sampling methods include trajectory parallelization, metadynamics and temperature-based methods. Trajectory parallelization methods use large groups of trajectories which are run in parallel, while metadynamics techniques use a history-dependent biasing potential that facilitates the escape of the system form free energy minima to recover unbiased kinetic information. Temperature-based methods use higher temperature to encourage the crossing of energy barriers. Other methods for the binding/unbinding problem include transition path sampling, smoothed potential molecular dynamics and accelerated molecular dynamics. Coupled with advances in computational hardware, these methods are beginning to reach timescales of pharmaceutical interest and have demonstrated promising quantitative agreement with experimental data. It is

therefore realistic to think that these techniques will allow the incorporation of kinetic properties earlier into the drug discovery pipeline (57).

In conclusion, the application of residence time in drug design and development offers the potential of a paradigm shift, from a static view in which drug efficacy is related to target affinity alone, to a dynamic view in which pharmacology is linked to the duration of target occupancy *in vivo* (52). In drug design, residence time must be carefully balanced, since a long residence time is not always a desirable effect in pharmacology, as it can lead to safety issues and side effects (58). Therefore, a deeper understanding of structure-kinetics relationships will help the development of better pharmacokinetic and pharmacodynamics models that can predict optimal dosage profiles, clearance rates, drug toxicity and therapeutic windows.

2. Aim of the research

GSK-3β is a point of convergence of multiple signaling pathways, and the deregulation of its activity plays a crucial role in the pathogenesis of diverse diseases such as diabetes, cancer, neurodegenerative disorders, chronic inflammatory and immunological diseases (5). GSK-3β has therefore become an attractive target for medicinal chemists to design and develop new inhibitors for pathologies that have currently many limitations in therapeutic treatment. Despite wide research for the past 20 years, kinase selectivity remains one of the main challenges in the design of GSK-3β ATP-competitive inhibitors. Recently, the design of allosteric modulators is gaining importance since this type of inhibitors allows increased selectivity and subtle modulation of the target. This is of particular importance for GSK-3β modulation as a therapeutic target, as its inhibitors should only inhibit the aberrant GSK-3β kinase activity implicated in pathological events.

The design of allosteric inhibitors is one of the strategies that can be applied to achieve kinase selectivity. A different possibility to gain kinase selectivity is the modulation of drug-target residence time during lead optimization. In fact, if a drug demonstrates a long residence time on its target and a shorter residence time on its off-targets, the selectivity is enhanced thus providing advantages for drug safety *in vivo* (50) (59).

Within this context, this research focuses on both approaches that are applied in two drug discovery projects. In the first project, GSK-3β is used as a reference system for the perspective validation of a novel computational method that has been developed within our group at the Italian Institute of Technology (IIT). This method, based on adiabatic bias molecular dynamics simulations, aims to simulate protein-ligand unbinding events and to rank series of compounds. To confirm a coherence between predicted and experimental k_{off}, a highly congeneric series of ATP-competitive GSK-3β inhibitors is considered, and protein-ligands k_{off} values are calculated through surface plasmon resonance (SPR) experiments. To provide a reliable configuration for computational simulations, X-ray structures of selected compounds in complex with GSK-3β are also generated. To achieve this aim, the first step is the setup of an efficient expression and purification protocol for

GSK-3β in eukaryotic cells. Protein expression and X-ray crystallography are carried out in collaboration with the Protein production lab at Elettra Synchrotron in Trieste.

The second project is focused on the discovery and characterization of GSK-3β allosteric inhibitors. To date, no crystal structure of GSK-3β bound to an allosteric inhibitor has been resolved yet, therefore a detailed structural knowledge and a mechanistic understanding of allosteric regulation is still lacking. Computational and biophysical techniques are employed to discover novel allosteric GSK-3β inhibitors and, in parallel, to further characterize allosteric modulators already described in the literature (46). In order to explore druggable binding sites on GSK-3β surface, a computational analysis with the novel computational algorithm Pocketron is performed at Italian Institute of Technology by our collaborators. Starting from this analysis, a Virtual Screening (VS) investigation on IIT compounds library is carried out targeting the most interesting pockets. The identification of compounds with allosteric properties is also pursued with a fragment-based approach (FBA) through NMR spectroscopy, by screening an in house library of fluorinated fragments.

In parallel, a couple of allosteric inhibitors described in the literature are also considered. In 2017, Gao and coworkers synthesized a series of benzothiazinones (BTOs) which acted as non-ATP competitive GSK-3β inhibitor and showed suppression efficacy on ovarian cancer (46). Biophysical and structural studies are carried out on two compounds of this series in order to further characterize their allosteric interaction with GSK-3β and to provide details on the mechanism of binding.

In summary, the main goals of the present research are:

- Optimization of an efficient protocol for the expression and purification of full-length GSK-3β, and biophysical characterization of the purified protein
- Validation, through surface plasmon resonance (SPR) experiments and X-ray crystallography, of a computational protocol able to simulate protein-ligand unbinding events based on adiabatic bias molecular dynamics (ABMD).

- Investigation of potential allosteric pockets on GSK-3β surface through computational algorithms and discovery of novel allosteric inhibitors via Virtual Screening, Nuclear Magnetic Resonance (NMR) spectroscopy and Microscale Thermophoresis (MST).
- Biophysical and X-ray structural characterization of known GSK-3β allosteric inhibitors

3. Materials and methods

3.1. GSK-3β expression in HEK293T cells

3.1.1. DH5α cells transformation

An aliquot of 50 µl of E.*coli* DH5α competent cells was thawed on ice and transformed with 25ng DNA plasmid of pEF1-N-CMV-TEV-GSK3β vector (kindly provided by Dr. O'Connell, University of Dublin). After incubation for 30 minutes on ice, cells were subjected to heat shock at 42°C for 20" and then 2' on ice. Cells were grown in SOC medium at 37°C for 1 hour at 225 rpm, and then plated on LB agar plates supplemented with 100 µg/mL ampicillin for the selection of transformed cells.

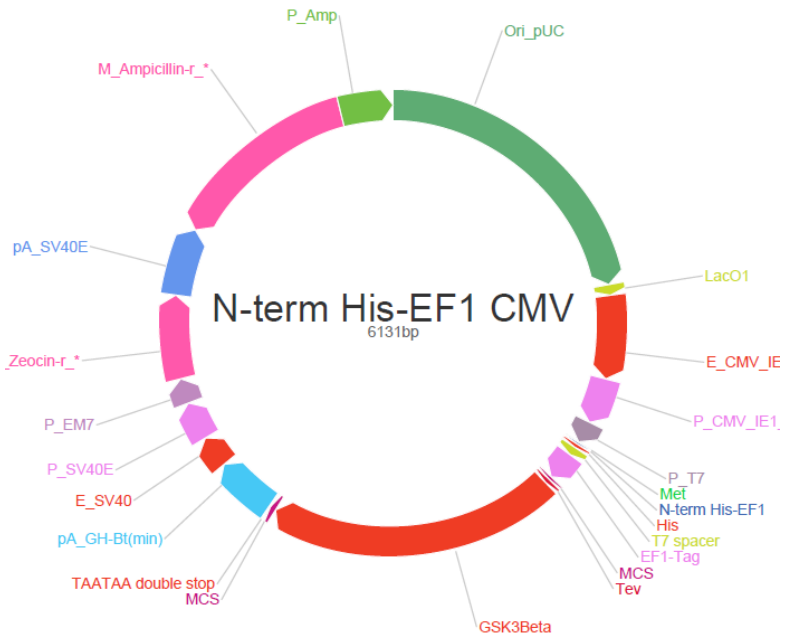


Figure 3. 1: pEF1-N-CMV-TEV-GSK3β plasmid. It encodes the EF1-tag at the N-terminus, CMV promoter, TEV cleavage size and ampicillin's resistance gene.

3.1.2. Plasmid DNA purification

A single colony of DH5α cells transformed with pEF1-N-CMV-TEV-GSK3β plasmid was grown at 37°C for 18 hours in 5mL of LB medium supplemented with ampicillin. Bacterial cells were harvested by centrifugation at 8000 rpm for 5'. Plasmid purification was carried out with Plasmid Maxi-Prep kit (Qiagen) following manufacturer's instructions.

3.1.3. HEK293T cells transfection

Human embryonic kidney cells (HEK293T) were transiently transfected with plasmid pEF1-N-CMV-TEV-GSK3β. HEK293T were cultivated in DMEM media supplemented with 10% fetal bovine serum (FBS), 1% penicillin/streptomycin and L-glutamine (2mM).

Cells were seeded at 800,000 cells/dish in 10-cm dishes 24 hours prior to transfection and transfected at 60-80% confluency using TransIT2020 transfection agent (Mirus). TransIT2020 reagent: DNA complex mix was generated using 5 µg plasmid DNA and 2.4 µL TransIT2020 per µg of DNA in 1 mL serum free media. Complex was incubated at room temperature for 20', and then added drop-wise to the cells. Cells were incubated for 24 hours at 37°C.

3.1.4. GSK-3β purification

Cells were washed with PBS and lysed in 0.75 mL high detergent lysis buffer (2 mM $CaCl_2$, 50mM Tris-HCl, 100 mM NaCl, 0.5% CHAPS (3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate hydrate), 0.1% CHS (cholesteryl hydrogen succinate), 1% DDM (n-Dodecyl β-D-maltoside), 30% glycerol and protease inhibitors (Roche)) and harvested using a cell scraper. Cell suspension was cleared by centrifugation at 10000xg for 20' at 4°C and pellet was discarded. Supernatant was pre-cleared using 50% slurry of unconjugated agarose resin on rotation at 4°C for 1

hour and then centrifuged at 1000xg for 1'. Supernatant was incubated with EF2-agarose resin (Konstant Biologics Ltd.) on rotation at 4°C overnight. Beads were collected by centrifugation at 5000xg for 1' at 4°C, washed with lysis buffer and protein was eluted in 80-200 µL of elution buffer (10 mM Hepes, 150 mM NaCl, 10 mM EDTA, pH 8.0).

3.2. GSK-3β expression in insect cells

3.2.1. Ligation Independent Cloning (LIC)

The DNA sequence of human GSK-3β full length 1-420 (UniProt ID: P49841) was PCR amplified from the cDNA template (GeneScript) with the following thermal protocol: 1' at 98°C, followed by 35 cycles of 10" at 98°C, 20" at 65°C, 30" at 72°C. Final extension was done for 2' at 72°C. PCR result was checked on 1% agarose gel. Primers sequences used for PCR are described in Table 3. 1.

	Sequence (in bold the GSK-3β specific sequence)	**T_m (gene specific) (°C)**	**GC content**
Fw	TACTTCCAATCCATG**TCAGGGCGGCCCAGAAC**	71.0	71%
Rev	TATCCACCTTTACTGTCA**GGTGGAGTTGGAAGCTGATGCAG**	71.0	57%

Table 3. 1. PCR primers sequences for cDNA amplification

Amplified cDNA was cloned in pFB-LIC-Bse vector (kindly provided by Dr. Opher Gileadi, SGC-Oxford) by Ligation Independent Cloning (LIC) (60).

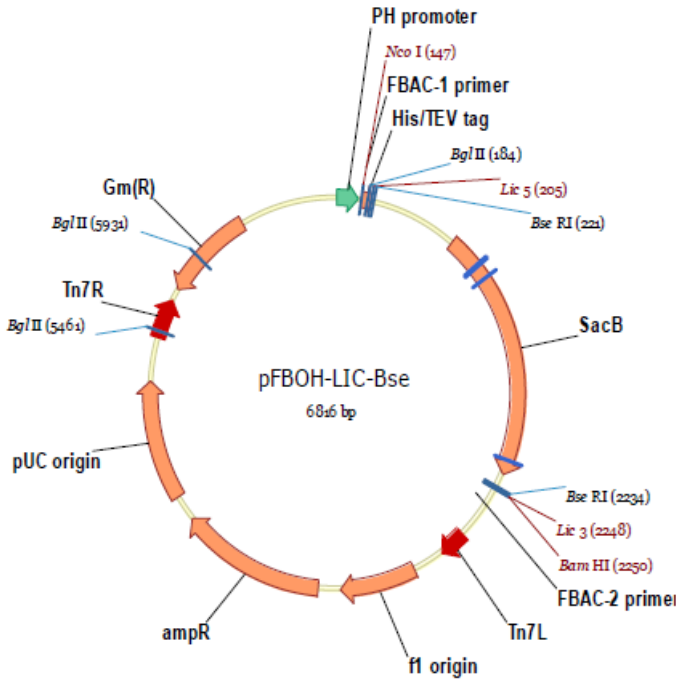


Figure 3. 2. pFB-LIC-Bse transfer vector. It encodes an N-terminal 6xHis-TEV cassette and includes sequences Tn7R and Tn7L that are the cleavage sites necessary for transposition and in vivo recombination into bacmid DNA. AmpR confers resistance to ampicillin and SacB sequence encodes for SacB protein (levansucrase) that converts sucrose into a toxic product, allowing negative selection of recombinant vectors on agar plates containing 5% sucrose.

pFB-LIC-Bse was linearized by BseRI enzyme upon incubation for 2 hours at 37°C and linearization was checked on 0.8% agarose gel. Digested vector and DNA insert were both treated with T4 DNA polymerase for 30' at 22°C and then for 20' at 75°C to inactivate the polymerase. Annealing of the DNA insert and the LIC vector was performed by mixing 0.02pmol of insert with 25-50 ng of vector and incubating at room temperature for 20'. DH5α competent cells were transformed with the vector by heat shock (30' on ice, 30" at 42°C, 2' on ice), and grown in SOC medium for 1h at 37°C. Transformed cells were selected on LB agar plate containing ampicillin and 5% sucrose, and incubated overnight at 37°C. A single colony was picked and grown in LB medium + ampicillin at 37°C overnight. DNA was purified using miniprep kit (Macherey-Nagel) following manufacturer's instructions, and sequence verified by Sanger sequencing. After sequence

verification, recombinant plasmid was transformed into E.*coli* DH10Bac competent cells by site-specific transposition to generate the recombinant bacmid DNA.

3.2.2. Bacmid generation and purification

DH10Bac competent cells were transformed with recombinant plasmid by heat shock (30' on ice, 45" at 42°C, 3' on ice). Transformed bacterial suspension was grown in SOC medium for 7-8 hours at 37°C under agitation to allow for transposition, diluted 1:50 in LB medium and spread on a LB agar bacmid plate (kanamycin 50 µg/mL, tetracyclin 10 µg/mL, gentamycin 7 µg/mL, IPTG 40 µg/mL, bluo-gal 100 µg/mL). Plate was incubated at 37°C protected from light for 48 hours or until the blue/white colonies could be differentiated. White colonies (positive clones containing the recombinant bacmid) were picked and restreaked on a new LB agar bacmid plate to verify the positiveness of the clone. A single white colony was inoculated in LB containing kanamycin 50 µg/mL, tetracyclin 10 µg/mL, gentamycin 7 µg/mL and grown overnight at 37°C.

Bacterial culture was centrifuged for 10' at 1500xg. Cells pellet was resuspended with 300 µL of resuspension buffer P1 (Qiagen). 300 µL of lysis buffer P2 were added and incubated for 5' at RT, then neutralized with 300 µL of buffer P3 and the suspension was centrifuged for 10' at 16000xg, RT. Centrifugation was repeated twice to ensure the complete removal of the precipitate. 100% isopropanol was added to the supernatant (to obtain a final concentration of 40%) and the mix was incubated on ice for 10' and centrifuged at 16000xg for 10' at 4°C. After centrifugation, supernatant was removed and 200 µL of 70% ethanol were added drop by drop to the pellet and centrifuged for 5' at 16000xg, RT. An additional wash of the pellet with 50 µL of 70% ethanol was done. Pellet was air-dryed under laminar flow hood and resuspended in 50 µL sterile water.

PCR on bacmid DNA was performed to confirm the transposition and verify the presence of the gene of interest using the M13 Forward primer (5'-GTTTTCCCAGTCACGAC-3') and M13 Reverse primer (5'-CAGGAAACAGCTATGAC-3'). Amplification of the target gene was done with initial

denaturation for 2’ at 95°C, followed by 30 cycles of 15” at 95°C, 15” at 55°C, 40” at 72°C. The reaction was terminated after a final extension for 10’ at 72°C.

3.2.3. FuGENE® Sf9 cells transfection

Sf9 cells (Expression Systems LLC, Davies - USA) were seeded in a six-well plate at $1.5x10^6$ cells/well in ESF-921™ medium (Expression Systems) and transfection was performed using FuGENE® HT reagent (Promega). Plate was incubated at 27°C and recombinant baculovirus was harvested 55 hours post-transfection (P0 stock). P0 was stored at 4°C protected from light and 2% fetal bovine serum (FBS) was added to stabilize the virus.

The high titer virus stock (P2) was generated by two rounds of amplification and used for protein expression. High Five (H5) cells (Expression Systems LLC, Davies - USA) were infected with P2 stock at an initial density of $1.5x10^6$ cells/mL in ESF-921™ medium, incubated at 27°C and harvested by centrifugation at 1000xg 72 hours after infection. After a wash step in PBS+10% glycerol, the pellet was stored at -80°C until purification.

3.2.4. Transfection optimization (PEI-MAX Sf9 cells transfection)

Transfection protocol was further reviewed and optimized to increase transfection efficiency and reduce multiple experimental steps. A new protocol based on a recent paper (61) was set up to allow the generation of a high titer P0 virus generation without the need to perform two rounds of amplification. Sf9 cells (Expression Systems LLC, Davies - USA) were seeded at $0.8x10^6$ cells/mL in a shaking flask and grown in ESF-921™ medium (Expression systems) for 3 hours at 130 rpm and 27°C. Transfection was performed using PEI-MAX solution by mixing 25 µg of bacmid diluted in 100 µL of PBS and 50 µL of PEI-MAX (1 mg/mL) and incubating for 30’ at RT. The mix was then added to Sf9 cells which were grown until signs of infection were evident (cell size increase; cell arrest). The high titer P0 virus was used for the infection of H5 cells at $1x10^6$ cells/mL in ESF-921™

medium, incubated at 21°C and harvested by centrifugation at 1000xg 96 hours after infection. After a wash step in PBS+10% glycerol, the pellet was stored at -80°C until purification.

3.2.5. GSK-3β purification

Cell pellets were resuspended in lysis buffer (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 10 mM imidazole, 5% glycerol, 5 mM $MgCl_2$, 1 mM DTT and 1x Complete protease inhibitors (Roche)) and lysed either by sonication at 50% amplitude (10" pulse, 20" pause on ice, total time 2 minutes), or by two rounds of homogenization (Emulsiflex – Avestin) at 1000 bar.

Lysate was incubated with Benzonase® (Merck-Millipore) for 10' at RT, and clarified by centrifugation for 1 hour at 30000xg, 4°C. The supernatant was loaded on equilibrated NiNTA agarose resin (Qiagen) on gentle rotation for 1 hour at 4°C. The flow-through was collected by mild centrifugation at 500xg for 15' and the NiNTA resin loaded on a polypropylene column (Bio-Rad). The column was washed with 20 CV of binding buffer (20 mM Tris-HCl pH 8.0, 0.5 M NaCl, 10 mM imidazole, 5% glycerol, 1 mM DTT) and eluted in elution buffer (20mM Tris-HCl pH 8.0, 0.5 M NaCl, 0.3 M imidazole, 5% glycerol, 1mM DTT).

Collected fractions were diluted 10-fold in Buffer A (20 mM Hepes pH 7.5, 5% glycerol, 1 mM DTT) to lower salt concentration and further purified by cation exchange chromatography on HiTrap SP column (GE Healthcare). Protein was eluted by applying a step salt gradient of Buffer B (20 mM Hepes pH 7.5, 1 M NaCl, 5% glycerol, 1 mM DTT). All protein molecular weights and phosphorylation states were confirmed by LC mass spectrometry. Purified protein was stored in small aliquots at -80°C in 20 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT.

3.3. Protein characterization

3.3.1. SDS-PAGE

SDS-PAGE was performed using alternatively either precast polyacrylamide gels (NuPAGE 4-12% BisTris Gel, Invitrogen) or hand casting polyacrylamide 12% gels. Loading buffer was added to protein samples prior to denaturation at 95°C for 5'. Precast gels were run in XCell SureLock Mini-Cell Electrophoresis System (Invitrogen) in MOPS SDS running buffer with a constant voltage of 70 mA for about 90'. Hand casting gels were prepared following Bio-Rad's instructions and run in Mini-PROTEAN Tetra Handcast chamber with a constant current of 200 V for about 1 hour.

3.3.2. Western Blot

SDS-PAGE gels were performed as described in section 3.3.1. Gel was blotted onto a 0.22 μm nitrocellulose membrane (Protran BA83, GE Healthcare) applying a constant voltage of 30 V for 2 hours in NuPage transfer buffer (ThermoFisher) with 10% isopropanol. Membrane was blocked for 30 minutes in blocking solution (5% no-fat milk-TBST) at room temperature, and probed overnight at 4°C with primary antibody diluted in 5% no-fat milk-TBST. Primary antibody was detected with a HRP-conjugated anti-rabbit secondary antibody (Abcam) diluted 1:10000 in 5% no-fat milk-TBST by incubation for 1 hour at room temperature. The membrane was detected with Amersham ECL Prime Western Blotting Detection Reagents (GE Healthcare) and exposed to ImageQuant LAS 4000 (GE Healthcare).

3.3.3. Thermal Shift Assay (TSA)

Thermal Shift Assay (TSA) was performed to test the stability of the protein both in the apo form and after incubation with inhibitors. For ATP-competitive inhibitors, 0.1 mg/mL GSK-3β was first incubated for 2 hours at room temperature with 3x molar excess of inhibitor. For allosteric inhibitors, the protein was incubated with 10x molar excess of AMP-PNP, 3x molar excess of inhibitor, 2 mM $MgCl_2$. After incubation, reaction mixtures were prepared in white 96-wells plate in 20 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT, 5x Sypro Orange (Invitrogen) at final volume of 20 µL. Plate was centrifuged at 200xg for 1' and assay was performed in a Real Time PCR instrument (CFX96 Touch, Bio-Rad) with the following protocol: 5' at 5°C and 0.3°C increment every minute up to 90°C. Data were analyzed in CFX Manager software (BioRad) and exported to GraphPad Prism.

3.3.4. Liquid chromatography-mass spectrometry (LC-MS/MS)

Proteolysis for LC-MS/MS was performed from in-gel digestion of bands excised from SDS-PAGE gel. After excision, gel pieces were destained in 25 mM NH_4HCO_3, 50% acetonitrile (ACN) and incubated at 56°C for 1 hour in reducing solution (10 mM DTT, 50 mM NH_4HCO_3). After incubation, reducing solution was removed and digestion buffer (50mM NH_4HCO_3) was added for 5' at room temperature. Alkylation was performed in the dark for 45' in alkylating solution (50 mM iodoacetamide, 25 mM NH_4HCO_3). Gel pieces were washed twice in digestion buffer and once in ACN, and dried under nitrogen stream. 0.5 µg/µL trypsin dissolved in digestion buffer was added to the sample and incubated at 37°C overnight. Tryptic peptides were extracted twice using a solution of 5% trifluoroacetic acid in 50% ACN. Extracted peptides were dried under vacuum.

Dried peptides from each gel slice were dissolved in 3% ACN with 0.1% formic acid (FA). The volume corresponding to 2 µg of peptides was injected on a NanoAcquity chromatographic system and analyzed with a TripleTof 5600+ mass spectrometer equipped with a NanoSpray III ion source. The trapping phase was performed on a 180 µm x 20 mm Acquity C18 column for 4 minutes at a flow rate of 4.0 µL/min (1% ACN + 0.1% FA). Peptides were separated using a PicoFrit C18 column

(75 µm x 25 cm, from NewObjective Inc., Woburn, MA, USA) and eluted at 300 nL/min with a 1 hour gradient of ACN in water (3% to 45%, both eluents were added with % FA). ACN content was increased at 90% in 5 minutes and kept for additional 5 minutes. The system was then re-equilibrated to 3% ACN for 18 min. Eluted peptides were analyzed in positive ion mode with ion spray voltage at 2550 V, spray gas 1 kept at 5, curtain gas at 30 and declustering potential at 80 V. Source temperature was set at 75 °C. Scan range was set to 400-1250 *m/z* for MS survey and to 100-1600 *m/z* for each MS/MS acquisition with 100 ms as accumulation time each.

3.3.5. Fluorescent gel shift electrophoresis activity assay

The protocol for the assay was adapted from Choi et al (62). 50 ng of GSK-3β were incubated with 4 mM ATP, 0.01 mM Pep(p) in buffer 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 0.1 mM EDTA, 2 mM DTT, 0.01% Brij 35, in a final reaction volume of 20 µL. Pep(p) is the primed peptide GSK-3β substrate [FAM]KEEPPSPPQ[pS]PR that is phosphorylated by GSK-3β on the serine residue.

The mix was incubated for 90 minutes at 30°C, and then loaded on 1.5% agarose gel applying a constant voltage of 50 V for 1 hour in Tris-Borate (TB) buffer. Fluorescence was acquired with SYBR green detection mode on ImageQuant LAS 4000 (GE Healthcare).

3.4. Biophysical methods

3.4.1. Agarose gel electrophoresis

Agarose gel electrophoresis is a method used to separate biomolecules such as nucleic acids and proteins, in which an electric field is applied in an agarose matrix. Agarose gel electrophoresis was performed using 0.8% or 1% agarose gel. Gel loading dye was added to the samples prior to loading DNA samples. Running of DNA samples was carried out applying a constant voltage of 80 V for 50

minutes in the horizontal electrophoresis apparatus (Mini-Sub® Cell GT Systems, BioRad) in TAE buffer supplemented with SYBR Safe DNA gel stain (Invitrogen).

3.4.2. Protein and DNA quantification

Protein absorbance was measured at 280nm using a UV-Visible Nanodrop2000 spectrophotometer (ThermoFisher) and protein concentration was calculated according to the Lambert-Beer law. Extinction molar coefficient ε was calculated using ProtParamTool (www.expasy.org/protparam).

	ε ($M^{-1}cm^{-1}$)	**Molecular Weight (Da)**	**Theoretical pI**
His-GSK-3β	39310	49222.93	8.80

Alternatively, protein concentration was measured with Bradford Protein Assay (Bio-Rad) using a BSA standard curve ranging from 1 mg/mL to 0.1 mg/mL concentrations. Absorbance was measured at 595nm on a Infinite200PRO TECAN plate reader.

Plasmid and bacmid DNA concentrations were obtained measuring the absorbance of the nucleic acids at 260 nm with Nanodrop2000.

3.4.3. Surface Plasmon Resonance (SPR)

Immobilization

SPR measurements were carried out on a Pioneer FE (Pall FortéBio) at 25°C. GSK-3β was immobilized on a HisCap SensorChip (Pall FortéBio), and then let equilibrate in immobilization buffer (20 mM Hepes pH 7.5, 150 mM NaCl, 0.01% Tween20).

The SensorChip surface was activated with 500 µM of $NiCl_2$ in immobilization buffer at a flow rate of 10 µL/min. After the activation of the surface, GSK-3β diluted in immobilization buffer at 7.5

µg/mL was immobilized by injecting 100 µL at a flow rate of 10 µL/min. The surface was stabilized by amine coupling with the injection of 7.5 µL of 0.1M NHS/0.4M EDC diluted 1:10 in H_2O at a flow rate of 15 µL/min, followed by an injection of 35 µL of 1.0 M ethanolamine pH 8.0 at a flow rate of 5 µL/min to block unreacted groups. The chip surface was cleaned from Ni^{2+} ions by injection of 60µl EDTA in immobilization buffer at a flow rate of 20 µL/min. The covalent coupling procedure resulted in a stable baseline. GSK-3β was immobilized alternatively on Flow cell (FC) 1 or FC 3, whereas FC 2 was used as reference. Typical immobilization levels ranged from 3500 to 4500 RU.

SPR Binding Assay

Binding experiments were performed in binding buffer (50 mM Tris-HCl pH 7.5, 250 mM NaCl, 0.01% Tween20) supplemented with 5% DMSO. All compounds were solubilized in 100% DMSO and then diluted in binding buffer by serial doubling. The top concentration for each compound was optimized to improve the accuracy of the steady-state and kinetic dissociation constants, K_Ds. In binding assays, a flow rate of 40 µL/min was set up. Compounds KDI, C44 and C50 did not require a regeneration step. For those compounds, association was measured for 5 minutes, and dissociation for up to 20 minutes depending on peculiar off-rates. Due to high off-rates, compounds L6Q, C22 and GR9 required a regeneration step. Association and dissociation were recorded for 3.5 and 2 minutes, respectively. An optimized regeneration solution, i.e., 15% DMSO in H_2O, was injected at a flow rate of 30 µL/min for 2 minutes to allow the complete dissociation of tight protein-ligand complexes.

Analysis of binding data

All data analysis and processing was performed using Pioneer Qdat Software (Pall FortéBio). Response (in RU) was recorded as a function of time. The equilibrium analysis was performed plotting the responses against the analyte concentration and fitting the data to a 1:1 binding isotherm. Kinetic curves were fitted to the Simple model of Qdat Software including bulk refractive index offsets. This parameter was particularly important to consider, since DMSO alters the refractive index. Binding analysis of all compounds injected across immobilized GSK-3β was replicated two

times on two different sensor chips with the same capacity. Although fitting data from multiple capacity surfaces provides more information about binding reactions, in most cases enough binding information is available from a single surface (63). K_D values obtained from equilibrium fit and kinetic fit agree if saturation in the experiment is achieved. A low difference between K_D values derived in both ways (dK_D) was used as an indicator to assess the quality of the experiments. For each compound, only the replicate that showed the lower dK_D was considered for k_{off} rates ranking.

3.4.4. Microscale Thermophoresis (MST)

MST analysis were performed on NT.115 Monolith (NanoTemper Technologies) with premium capillaries. GSK-3β was labeled with NT-647 red dye (NanoTemper Technologies) that binds the hexahistidine tail of the protein, following the manufacturer's protocol. All measurements were carried out at 25°C in MST buffer 20 mM Hepes pH 7.5, 150 mM NaCl, 2 mM $MgCl_2$, 0.01% Tween20 additioned with 5% DMSO to increase the solubility of the compounds. For binding check experiments, protein was diluted to 5 nM and incubated for 30' at room temperature with 500 μM AMP-PNP. After incubation, the mixture was centrifuged for 10' at 12000 rpm and incubated with 100 μM of each compound for 30' at room temperature.

For binding affinity experiments, the same procedure was applied but serial dilutions of each compound were prepared in MST buffer. Experiments were analyzed with MO.Control and MO.Affinity analysis software (NanoTemper Technologies).

3.4.5. Nuclear Magnetic Resonance (NMR)

NMR experiments were recorded at 25°C with a Bruker FT NMR Avance III 600-MHz equipped with a 5-mm CryoProbe QCI 1H/19F-13C/15N-D-Z quadruple resonance, a shielded z-gradient coil, and the automatic sample changer SampleJet NMR system with temperature control.

Activity assay

20 nM GSK-3β was incubated with 60 μM of the pre-phosphorylated fluorinated substrate Mimotope 2160204 (KRREILSRRP(p)SYRF(3-CF3)-OH) and 182 μM ATP in buffer 60 mM Hepes pH 7.5, 8% D_2O, 10 mM $MgCl_2$, 25 mM NaCl, 0.003% Triton, 50 μM EDTA, 2 mM TCEPT in a final reaction volume of 500 μL. 1D ^{19}F NMR signals were registered every twenty minutes for a total acquisition time of two hours. The reaction was quenched with the addition of 20 μM staurosporine.

For the calculation of the Michaelis-Menten constant K_M of ATP, GSK-3β pool 1 was incubated with 300 μM of Mimo2160204 and increasing concentrations of ATP. For the calculation of the K_M of the fluorinated substrate, GSK-3β was incubated with 400 μM ATP and increasing concentrations of Mimo2160204. To determine IC_{50} of staurosporine, 10 nM GSK-3β pool 1 was incubated with 30 μM of fluorinated substrate and 64 μM ATP in 60 mM HEPES, pH 7.5, 8% D_2O, 10 mM $MgCl_2$, 25 mM NaCl, 0.003% Triton in a final volume of 500 μL. Reaction was quenched after 90'. The inhibition constant (K_i) of staurosporine was calculated from the observed IC_{50} through the following equation: $K_i = IC_{50}/1+([ATP]/K_M)$.

Fragment screening

For the NMR Fragments screening a monodimensional ^{19}F NMR spectrum and a R_2 filter experiment were recorded, for a total of 64 scans with a repetition time of 5s and proton decoupling during the acquisition period. A spin-echo scheme with total $\tau = 0.2$ s was used for the R_2 filter experiments. Data were multiplied by an exponential function corresponding to a line broadening of 1 Hz prior to Fourier transformation. $CFCl_3$ was the reference standard in the ^{19}F spectra.

A home made local Environment of Fluorine (LEF) library constituted by 350 fluorinated compounds was screened after incubation of GSK-3β for 10 minutes with 1 mM AMP-PNP in order to saturate the ATP-binding pocket. Fragments were screened in mixtures of 20-25 compounds each at 40 μM concentration.

3.4.6. Crystallization

Sample preparation

Crystallization trials of GSK-3β were conducted using Hampton Research PEG/Ion Screen I and II in a 96-well plate dispensing 100 nL protein and 100 nL reservoir with Mosquito® TTP (Labtech). Optimizations of PEG3350 concentration, buffer and pH were carried out to achieve well diffracting crystals. Co-crystals of phosphorylated (pTyr216) his-tagged GSK-3β with compounds C44 or C22 were obtained by sitting drop method mixing 0.1 µL of protein-ligand solution with 0.1 µL reservoir (C44) or 0.5 µL of protein-ligand solution with 0.5 µL reservoir (C22) and equilibrating against 80 µL of precipitant solution at 20°C. Prior to crystallization each protein-ligand solution was obtained adding 600 µM compound to a solution containing 200 µM GSK-3β in 20 mM HEPES pH 7.5, 150 mM NaCl, 20 mM DTT, 5% glycerol (w/v), and incubating at room temperature for two hours. The protein-ligand solution was then mixed in the sitting drop with a reservoir solution containing 8% tacsimate pH 7.0 and 20% PEG3350 for compound C22 and 0.1M Bis-Tris pH 6.5, 2% tacsimate pH 6.0 and 20% PEG3350 for compound C44. Single crystals appeared within 4 days.

Phosphorylated (pTyr216) his-tagged GSK-3β crystals in complex with C50 were obtained by sitting drop and subsequent soaking. Protein crystals were grown at 20°C by mixing 0.1 µL of 150 µM GSK-3β, 20 mM HEPES pH 7.5, 150 mM NaCl, 20 mM DTT, 5% glycerol, 1 mM AMP-PNP, 2 mM $MgCl_2$ with 0.1 µL precipitant solution containing 0.2 M KF and 22% PEG3350. Resulting crystals were soaked for two hours in 1.2 mM C50 in 0.2M KF and 22% PEG3350.

For allosteric inhibitors 20g and 20f crystallization, GSK-3β in the phosphorylated (pTyr216), unphosphorylated and unphosphorylated without the his tag form was concentrated at 120 µM and was incubated with 2 mM MgCl2, 20 mM DTT, 1.2 mM AMP-PNP for 2 hours at room temperature. Protein solution was mixed with mother liquor and incubated at 20°C until crystals appeared. Allosteric inhibitors were added directly on the drop at 1.2 mM final concentration and incubated for a minimum of 4 hours prior to freezing.

Crystals were cryo-protected in mother liquor with 20-25% glycerol and 3X molar excess of inhibitor, and flash-frozen in liquid nitrogen prior to data collection.

Data collection and structure determination

X-ray data were collected at beamline XRD1 of Elettra Synchrotron, Trieste, Italy. Crystals corresponding to ligand C22 and C44 diffracted to 2.2 and 2.35 Å respectively, while ligand C50 diffracted to 3.2 Å. Integration of reflection files was performed using XDS (64) and integrated data was scaled using AIMLESS (65) in CCP4 suite (66). Phasing was performed by molecular replacement using phaser (67). Previously published GSK-3β structure (PDB ID: 4ACD) was used as initial search model in phaser. Structures were refined using phenix.refine (68) and model optimization and editing were performed in Coot (69). In the case of data corresponding to ligand C50, refinement was performed using low resolution refinement in REFMAC (70), wherein the restrains were generated using PDB model 4ACD in proSMART.

3.5. Computational methods

3.5.1. Adiabatic bias molecular dynamics simulations

Crystal structures of GSK-3β in complex with L6Q, GR9 and KDI (35) were retrieved from Protein Data Bank (PDB IDs 4ACG, 4ACD and 4ACH). For ligands C22, C44 and C50 experimental protein-ligand complex were not available, so new crystal structures were generated (PDB IDs 6HK4, 6HK3, 6HK7). Protonation state of all inhibitors was assigned at pH 7.0 with Molecular Discovery MoKa software (71). GSK-3β was prepared with Protein Preparation wizard in Maestro, and all crystallographic water molecules were deleted except the ones located in the internal ATP-binding site of complexes with ligands L6Q and GR9. All systems were prepared with BiKi Life Sciences 1.3 Software Package (72). MD simulations were performed using Gromacs 4.6.1 (73). The protein's structure was left free of restraints. To prioritize the series of congeneric compounds in terms of residence time, a statistics of 20 independent replicas of 30ns each was collected for each GSK-3β complex running on the Linux cluster Marconi at CINECA. Analysis of the simulations were

performed with BiKi Life Sciences 1.3 Software Package. Bootstrap analysis was performed to define robust mean estimations (63).

3.5.2. Pocketron analysis

The automated algorithm Pocketron (74) was used to examine protein formation, dynamics and allosteric communication with molecular dynamics (MD) simulations. Pocketron is a module implemented in the BiKi Life Sciences software suite (www.bikitechnologies.com) (72), which detects the formation and spatiotemporal evolution of all protein pockets, providing a detailed analysis of the structure-dynamics-function relationship from MD simulations. NanoShaper 0.7 was used as a preliminary tool to detect the pockets on GSK-3β surface on individual frames, with particular attention to those that could be druggable sites. Pocket dynamics was characterized along the molecular dynamics trajectory through the assignment of a unique identifier for every pocket. The spatiotemporal evolution of every pocket was monitored over the entire MD trajectory. MD simulations were 1 μs long and were run on GSK-3β PDB structure 1J1C.

3.5.3. Virtual Screening

Virtual screening on GSK-3β was performed on allosteric cavity 5 identified in the work by Palomo and colleagues (36). Reference crystal structures of GSK-3β were downloaded from the Protein Data Bank. Chains A of structures 4NM5, 3IB4 and chain B of 1J1C were used as reference for cavity 5. Six virtual screenings were performed in total. All structures were treated with Maestro Protein Preparation Wizard tool (Schrödinger) (75) and amino acids side chains were reconstructed. Water molecules and ions were removed, and an exhaustive sampling of the orientations of groups was performed. Finally, the protein structures were refined to relieve steric clashes with a restrained minimization with the OPLS2005 force field.

IIT library of compounds was prepared with the LigPrep tool of the Schrödinger suite. All ligands were docked with Glide SP (76) by centering the grid on GSK-3β residue Ser236, which was chosen on the basis of the druggability index calculated by SiteMap (77). Both grid generation and docking calculations were performed with the default settings. The compounds were selected by visual inspection and clustered based on chemical similarity.

4. Results and discussion

4.1. GSK-3β expression in HEK293T cells

GSK-3β was initially expressed in human embryonic kidney cells (HEK293T) exploiting a novel affinity tag technology developed by the University College Dublin and described by Dunning and coworkers (78) in a study investigating affinity interaction between amyloid β peptide Aβ42 and GSK-3. Mammalian cells were the expression system of choice since they can perform complex post-translational modifications (such as phosphorylation) which play a key role in the regulation of GSK-3β activity, as summarized in the introduction section.

The affinity tag system relies on the calcium-dependent interaction between two EF hand domains (EF1 and EF2) of human calbindin D9k. HEK293T were transiently transfected with pEF1-N-CMV-TEV-GSK3β vector, which incorporates EF1 expression hand domain in frame with GSK-3β. EF2 hand domain is conjugated with an affinity agarose matrix, which binds to EF1 in the presence of calcium. The transfection protocol was optimized by varying the ratio between plasmid DNA and transfection reagent (Figure 4. 1). Western blot analysis showed that a ratio of 1:2 between plasmid DNA and transfection reagent provided the highest level of protein expression.

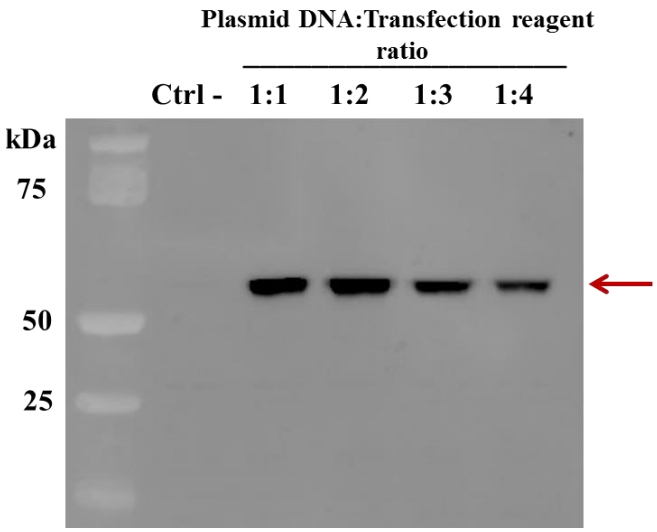


Figure 4. 1. Western blot analysis of transfection optimization. Different ratio between plasmid DNA and transfection reagent using primary antibody anti-calbindin D9k 1:1000 are shown.

After transfection, HEK293T cells were harvested with a cell scraper and lysed with a high detergent lysis buffer. Lysate was first incubated with an unconjugated agarose resin, then with EF2-conjugated resin. Protein was eluted with a buffer containing the chelating agent EDTA (10 mM Hepes, 150 mM NaCl, 10 mM EDTA, pH 8.0).

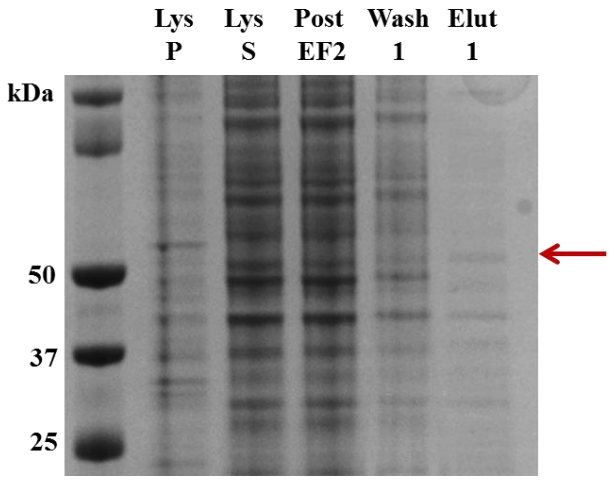


Figure 4. 2. SDS-PAGE analysis of EF1-GSK-3β expression in HEK293T cells. Lys P: pellet of cell lysate. Lys S: supernatant of cell lysate. Post EF2: flow through after incubation with EF2-conjugated resin. Wash 1: fraction collected after first beads wash with lysis buffer. Elute 1: first eluted fraction. Red arrow indicates EF1-GSK-3β protein.

Fractions corresponding to the expression and purification of EF1-GSK-3β were analyzed by SDS-PAGE (Figure 4. 2). A band at ~ 53 kDa, corresponding to EF1-GSK-3β protein was detected in most fractions, including the flow-through after incubation with EF2-conjugated resin, indicating that EF1-EF2 binding was probably not completely effective. In the eluted fraction, only a small amount of recombinant protein was present. Overall, the expression system was not suitable to achieve the desired purity.

Western blot analysis were also performed to further characterize purification fractions (Figure 4. 3). Monoclonal antibody anti-calbindin D9k (D-5) was used to identify recombinant EF1-GSK-3β, as it recognizes amino acids 1-60 mapped at the N-terminus of human calbindin D9k where EF1 hand domain is located.

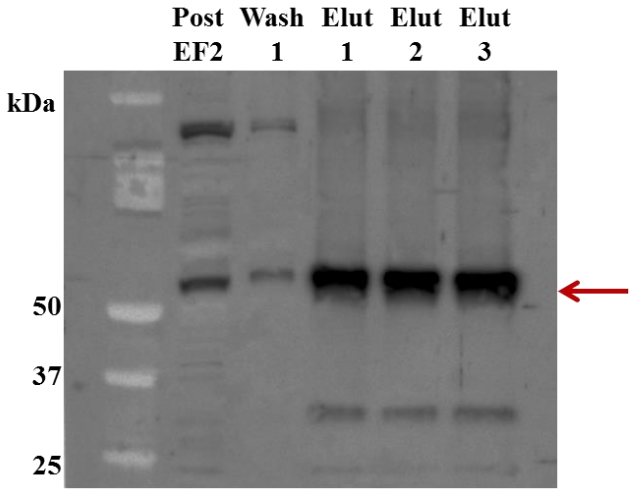


Figure 4. 3. Western blot analysis using primary antibody anti-calbindin D9k 1:1000. Post EF2: flow through after incubation with EF2-conjugated resin. Wash 1: fraction collected after first beads wash with lysis buffer. Elut 1-3: eluted fractions. The red arrow indicates EF1-GSK-3β protein.

EF1-GSK-3β was found in all fractions, but most of the protein was eluted in the final step of purification. Purification provided a total yield of 16 µg of protein, starting from 800000 seeded HEK293T cells. Despite the innovative approach, protein expression in mammalian cells with EF1 affinity tag provided an inadequate protein yield and purity for subsequent experimental purposes. A series of limitation of this method were highlighted, therefore instead of scaling up this protocol we

decided to employ a different method that could still provide post-translational modifications (such as phosphorylation), that play a key role in the regulation of GSK-3β activity.

4.2. GSK-3β expression in insect cells

The baculovirus expression system was the method of choice for the expression of GSK-3β, since it offers many advantages over traditional systems: the phosphorylation pattern is very similar to mammalian cells; it is easy to scale-up, as insect cells can be grown in suspension cultures in classical Erlenmeyer flasks under orbital shaking allowing thus large-scale protein production. Baculovirus-mediated protein expression in insect cells can provide high yields of recombinant protein, which is essential to ensure a proper drug discovery approach based on 3D structure determination and SAR analysis.

4.2.1. Bacmid generation

The first step for bacmid generation was the amplification of GSK-3β full-length sequence from the cDNA template. PCR result was controlled on agarose gel (Figure 4. 4), in which a band at 1260 bp (corresponding to the full-length coding sequence) was detected.

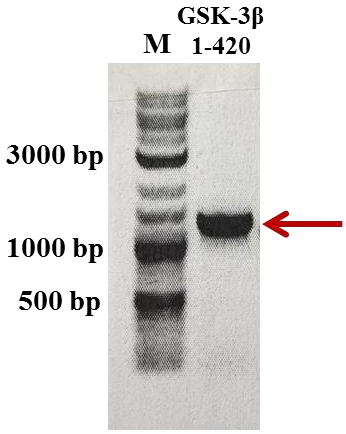


Figure 4. 4. 1% agarose gel of GSK-3β DNA sequence amplification

Amplified cDNA was then cloned in a pFB-LIC-Bse vector by Ligation Independent Cloning (LIC), which allows molecular cloning without the use of restriction endonucleases or DNA ligases. pFB-LIC-Bse is derived from the pFastBac vector by introducing LIC sites. In this method, primers used for PCR include 14-bases extensions identical to the sequences at the cloning sites in the vector. Both vector and insert are treated with T4 DNA polymerase, exposing complementary 14-nt single-stranded tails. The vector and insert are then annealed and introduced into E.*coli* DH5α cells (79) as described in the Material and methods section.

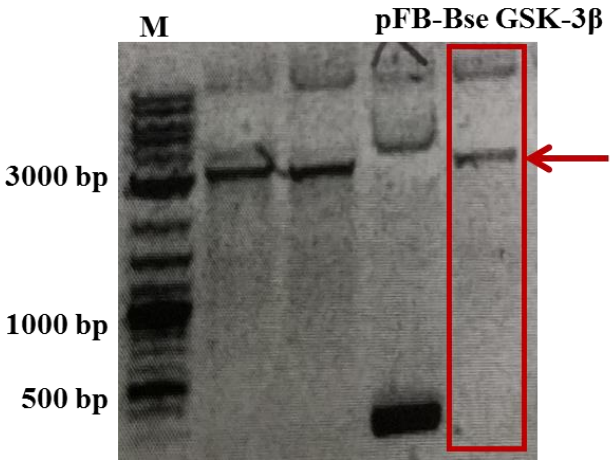


Figure 4. 5. 1% agarose gel confirming the correct insertion of the gene of interest in the recombinant bacmid

Transformed DH5α cells were selected and DNA was purified by miniprep kit. The recombinant plasmid was transformed into E.*coli* DH10Bac in order to generate the recombinant bacmid DNA. Agarose gel in Figure 4. 5 shows the result of colony screening by PCR. To verify the insertion of the gene of interest in the recombinant bacmid, M13 Forward and M13 Reverse primers were used, exploiting the M13 priming sites flanking the mini-attTn7 site within the lacZα complementation region in the bacmid. The insertion of the gene into the plasmid resulted in a PCR product of ~ 2400 bp plus the size of the insert (1260 bp), which was detected in the agarose gel. This result confirmed the correct transposition and insertion of the gene in the recombinant bacmid.

4.2.2. Small scale expression

After baculovirus generation by subsequent virus amplification rounds in insect cells, a small scale expression study was performed to identify the proper virus multiplicity of infection (MOI), cell line (Sf9 or H5) and harvesting time point (48h or 72h) to optimize expression conditions and obtain information regarding scale-up possibilities. In fact, these factors can significantly influence the expression levels of a heterologous gene.

Cells were infected with increasing viral loads and harvested after 48 or 72 hours. The infection procedure, incubation, harvesting and affinity purification were performed as described in Materials and methods section.

SDS-PAGE analysis of test expression (Figure 4. 6) allowed the selection of the optimal parameters. The highest protein yield was reached in H5 cells with MOI = 3 (corresponding to 33.3 µL virus/10^6 cells) after 48h of incubation, conditions that were chosen for scaling up protein expression.

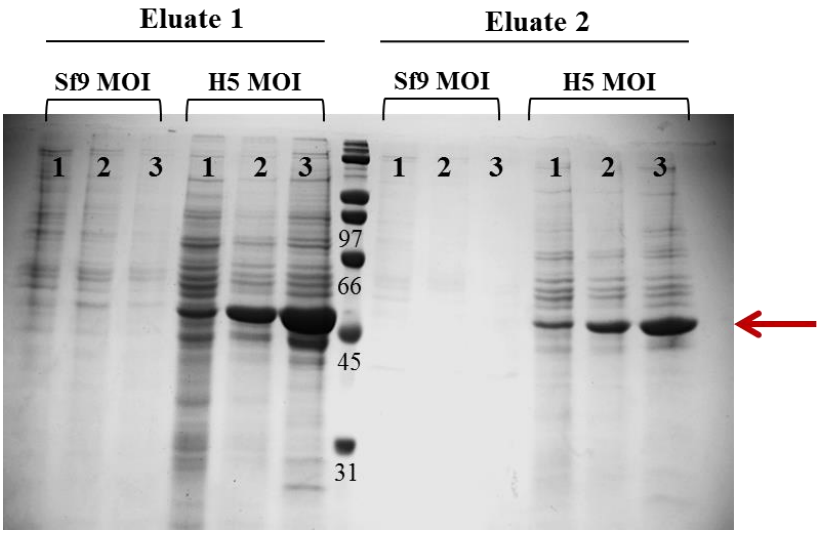


Figure 4. 6. SDS-PAGE analysis of GSK-3β test expression in Sf9 and H5 cells after 48h. MOI 1-3: multiplicity of infection

4.2.3. GSK-3β purification

The protocol for GSK-3β purification in H5 cells was optimized with two chromatographic steps: nickel-charged affinity resin (Ni-NTA agarose) followed by cation-exchange chromatography (HiTrap SP HP column).

Ni-NTA affinity resin

Ni-NTA agarose affinity resin was used to purify GSK-3β exploiting its high affinity and selectivity for the polyhistidine (6xHis) tag at the N-terminus. After mechanical cell lysis and incubation with benzonase to decrease viscosity due to DNA presence, supernatant was shortly incubated with Ni-NTA and the resin was transferred to a polypropylene column. Protein bound to the resin was eluted by competition with imidazole. SDS-PAGE gel (Figure 4. 7) of the Ni-NTA purification revealed an intense band in the eluate column at 49 kDa, corresponding to the expected molecular weight of GSK-3β. Only small fractions of the protein were lost during the different purification steps, indicating an efficient procedure.

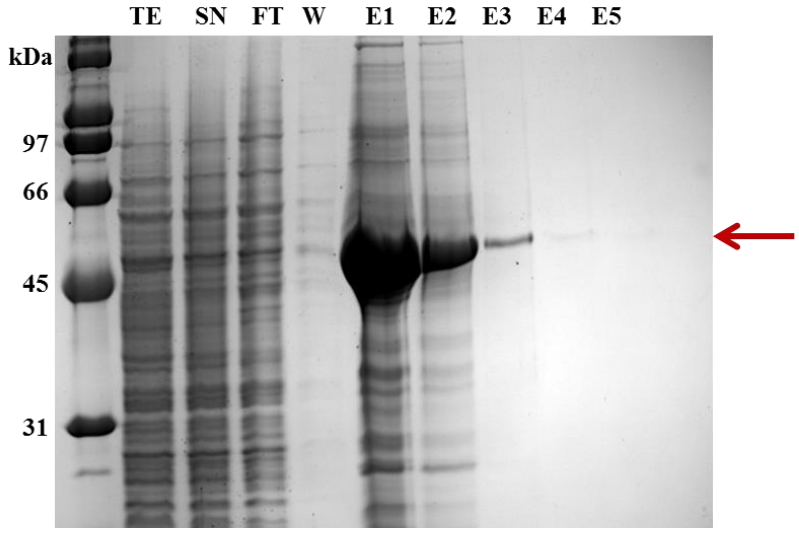


Figure 4. 7. SDS-PAGE analysis of GSK-3β NiNTA purification. TE: total extract. SN: supernatant after cell lysis. FT: flow through. W: wash. E1-E5: eluates

Cation-exchange chromatography

Pooled fractions eluted from Ni-NTA resin (diluted in order to reduce salt concentration) were loaded onto a HiTrap SP column. Cation-exchange chromatography was the method of choice, as the isoelectric point of GSK-3β is 8.8. The method uses a negatively charged ion exchange resin with high affinity for molecules having net positive surface charge.

The elution, initially performed by setting a linear gradient from 30 mM to 1 M NaCl, generated a chromatogram with one asymmetric main peak (Figure 4. 8). The chromatogram displayed in the picture refers to a purification of 0.75E9 of H5 cells. Absorbance at 280 nm (A_{280}) is plotted against the elution volume.

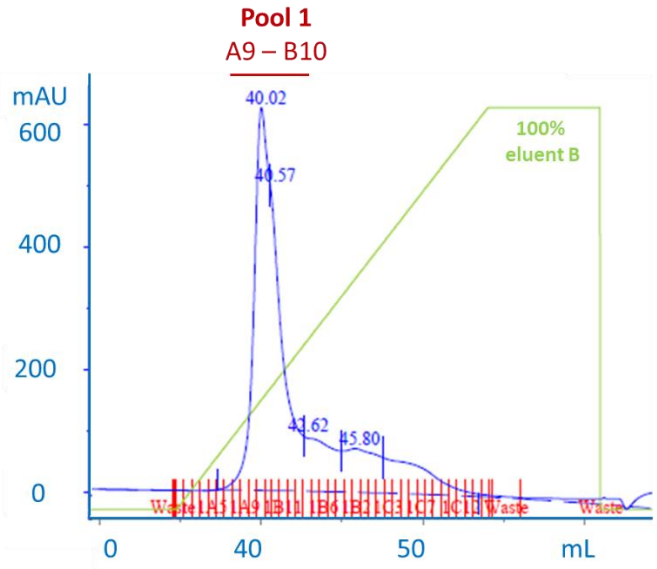


Figure 4. 8. Chromatographic profile of the linear gradient HiTrap SP column

Fractions corresponding to the peak eluted from the cation exchange column were analyzed by SDS-PAGE. The gel of the eluted fractions showed one main band corresponding to GSK-3β. A second band at slightly lower molecular weight, presumably corresponding to the peak's shoulder in the chromatogram, was also detected (Figure 4. 9). Overall, the fractions mainly contained pure GSK-3β.

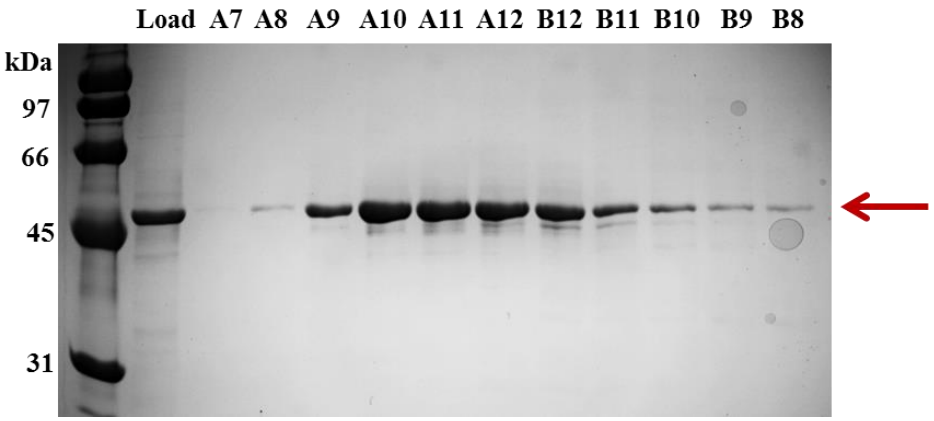


Figure 4. 9. SDS-PAGE analysis of linear gradient cation exchange chromatography

The final protein yield, quantified with Bradford Assay, was 2.3 mg per $1x10^9$ of H5 cells.

The main peak's shoulder present in the chromatogram could correspond to a second GSK-3β phosphorylation state. An example of separation of the GSK-3β phosphorylated and nonphosphorylated forms with cation exchange chromatography had already been reported in the literature (18). In order to overcome the SDS-PAGE gel double band issue, we applied a salt step gradient to the HiTrap column rather than a linear gradient. We were successfully able to elute GSK-3β in three separate peaks (Figure 4. 10). All three peaks contained GSK-3β, as shown by SDS-PAGE and Western blot analysis (Figure 4. 11).

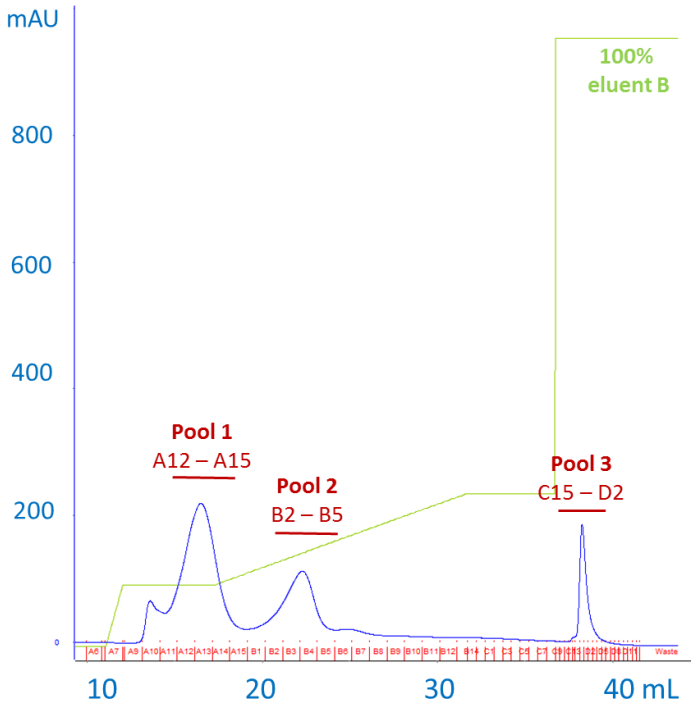


Figure 4. 10. Chromatographic profile of the step gradient HiTrap SP column

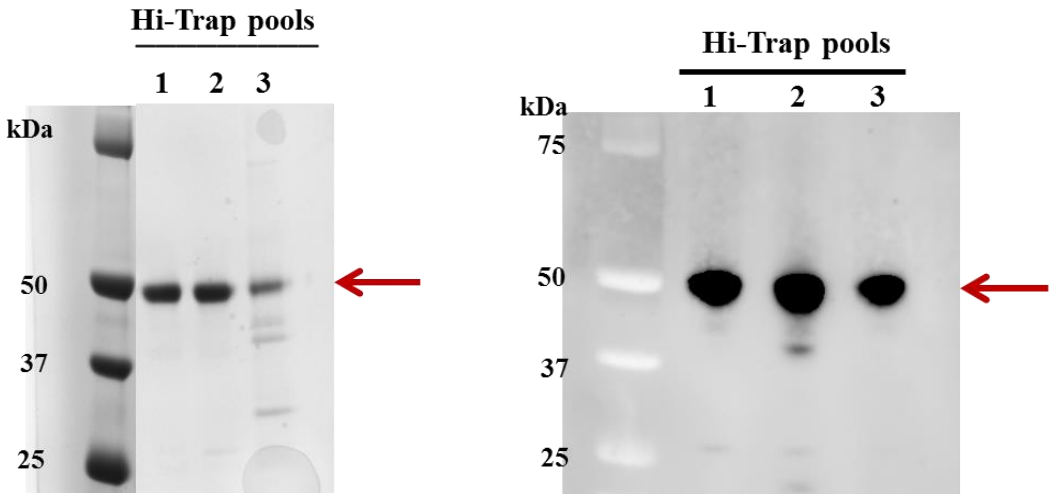


Figure 4. 11. Left: SDS-PAGE analysis of step gradient cation exchange chromatography pools. Right: Western Blot analysis with primary antibody anti-GSK3β 1:10000

In order to characterize the purified pools of GSK-3β, the protein was further analyzed with NMR, LC-MS/MS and TSA.

4.3. GSK-3β characterization

4.3.1. Characterization of GSK-3β activity

NMR

The characterization of purified GSK-3β activity was carried out with a sensitive NMR method named 3-FABS (three fluorine atoms for biochemical screening) that allows to set up functional biochemical screenings. This technique requires the labeling of the kinase substrate with a CF_3 moiety. The enzymatic reaction is performed with the CF_3 labeled substrate and fluorine NMR spectroscopy is used as the method of detection. The reaction is quenched after an established delay that depends on the enzyme, cofactors and substrate concentrations (80). In our case, signal was recorded every twenty minutes and the reaction was quenched by adding the strong GSK-3β inhibitor staurosporine. The substrate of choice was the mimotope Mimo2160204 (Lys-Arg-Arg-Glu-Ile-Leu-Ser-Arg-Arg-Pro-pSer-Tyr-Arg-Phe(3-CF3)-OH). R2 filter fluorine NMR experiments were then used to monitor the starting and enzymatically modified substrate.

Figure 4. 12 shows the ^{19}F NMR spectra recorded for the three main cation-exchange GSK-3β elution pools (namely, pool 1, pool 2 and pool 3). ^{19}F NMR spectra highlighted that the speed of the enzymatic reaction for the three pools was significantly different. Pool 1, which eluted first from the HiTrap column, resulted to be the most active. Pool 2 displayed a slower activity profile, while pool 3 did not show any activity.

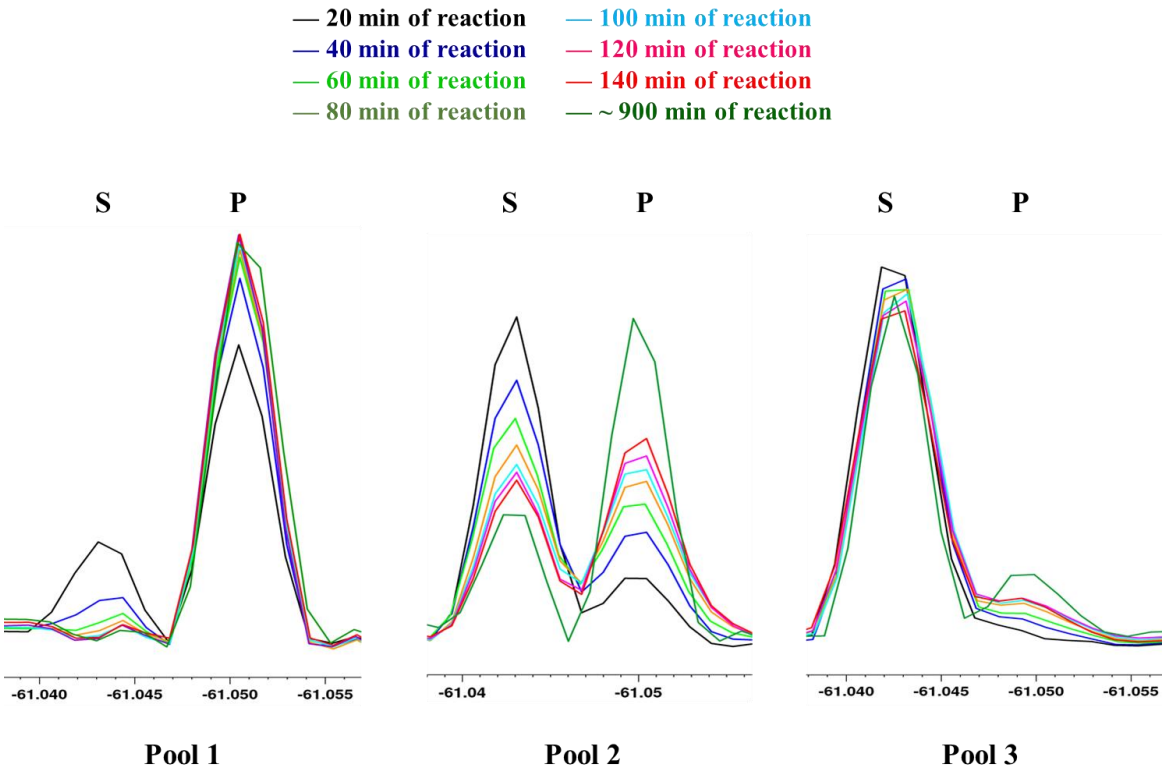


Figure 4. 12. ^{19}F NMR spectra for the three main GSK-3β cation-exchange elution pools. Chemical shifts are referred to trifluoroacetic acid. S and P are the substrate and the product, respectively.

Since pool 1 displayed to be the most active, this fraction was used for further NMR experiments, such as the calculation of K_M of ATP and fluorinated substrate and the measurement of IC_{50} for staurosporine binding.

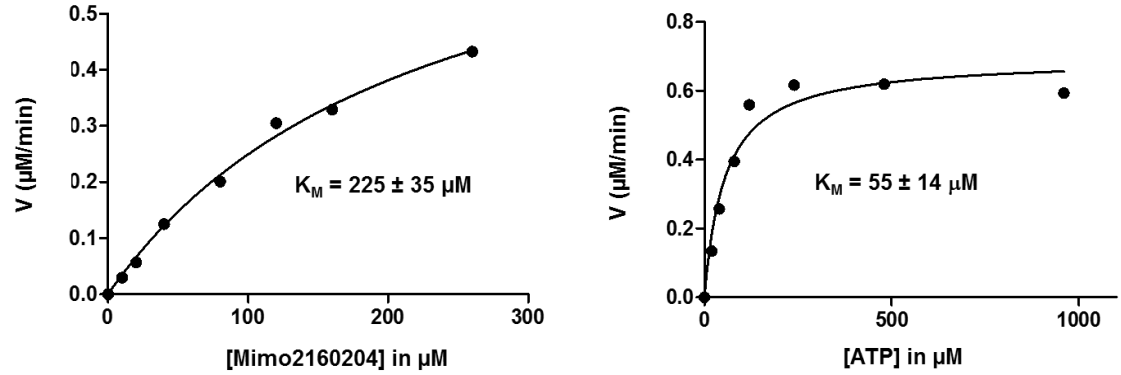


Figure 4. 13. K_M measurement of fluorinated substrate (left) and ATP (right) for GSK-3β

Figure 4. 13 shows K_M measurements of fluorinated substrate and ATP for GSK-3β. Each plot represents the speed (V) of the enzymatic reaction (^{19}F signal integral of the product divided by the incubation time) as a function of substrate or ATP concentration. K_M measurement of ATP was performed at different ATP concentrations with GSK-3β and substrate at fixed concentrations. The best fit of the experimental data gave a K_M for the substrate of 225 ± 35 μM and a K_M for ATP of 55 ± 14 μM. The latter value is coherent with literature data (34).

Experiments performed at different inhibitor concentrations allowed the determination of the IC_{50} (i.e., concentration of the inhibitor at which 50% of the reaction is achieved) of staurosporine as shown in Figure 4. 14. The product concentration (measured with integration of the ^{19}F NMR product signal) is plotted as a function of staurosporine concentration. According to these experimental conditions, the measured IC_{50} of 16.3 nM corresponds to a K_i of 7.4 nM. In the absence of allosteric effects and assuming a purely competitive mechanisms, the inhibition constant (K_i) is related to the IC_{50} through the following equation: $K_i = IC_{50}/1+([ATP]/K_M)$. The value obtained is remarkably close to the value reported in the literature (30).

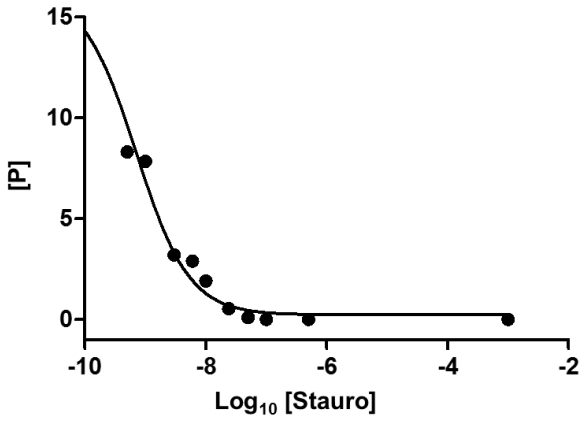


Figure 4. 14. IC_{50} measurement of staurosporine for GSK-3β

Fluorescent gel shift electrophoresis

Fluorescent gel shift electrophoresis activity assay was employed as a rapid way to qualitatively assess the activity of GSK-3β pools after HiTrap purification. The method was first developed by Choi and coworkers (62) as a way to detect GSK-3 activity in mouse sperm using a fluorescent

agarose gel mobility shift assay. The assay employs the dye-tethered and primed peptide substrate for GSK-3 [FAM]KEEPPSPPQ[pS]PR, which is known to be consecutively controlled by MAPK and GSK-3. The principle of the assay relies on the fact that peptide electrophoresis shows the charge difference in peptide status, causing the phosphorylated form to shift down toward the positive electrode. Upon phosphorylation by GSK-3, a distinct mobility shift in the fluorescent bands of agarose gel is observed, due to the increased negative charge of the phosphate groups (80 Da) of the primed peptide (PEP (p)). As shown in Figure 4. 15, a band shift was observed in pool 1 and pool 2, confirming the phosphorylation of the primed peptide in the presence of ATP and therefore displaying catalytic activity. Contrariwise, no band shift of in pool 3 was detected, indicating the lack of catalytic activity. Commercial GSK-3β was used as a positive control.

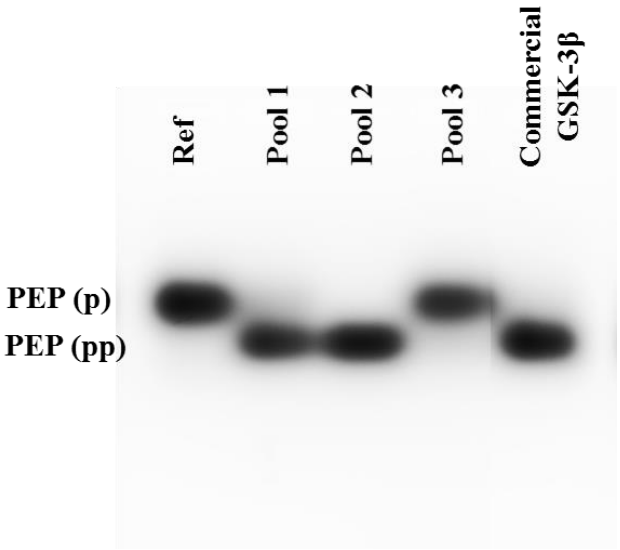


Figure 4. 15. Agarose gel mobility shift assay of GSK-3β pools

These observations were in line with NMR activity assay, and further confirmed the different catalytic activity among cation exchange chromatography GSK-3β pools.

4.3.2. Characterization of GSK-3β phosphorylation pattern

To investigate the structural basis of the differences observed in GSK-3β catalytic activity, a mapping of the phosphorylation sites was carried out for each pool obtained from the HiTrap purification by liquid chromatography/tandem mass spectrometry (LC-MS/MS). It has been shown that the activity of GSK-3β is regulated by the phosphorylation of residues Tyr216 (which increases its catalytic activity), and Ser9 (which inactivates the protein) (10). We therefore explored how the phosphorylation pattern of the purified protein could influence the behavior observed in the previous experiments. For each protein pool, the extracted ion chromatogram (XIC) corresponding to both the non-modified peptide and the phosphorylated peptide was searched. Table 4. 1 reports the peptides sequences searched for each phosphorylation site, together with the mass-to-charge ratio (*m/z*) of the naïve and the phosphorylated form.

Phosphorylation site	Peptide	Naïve (*m/z*)	Phos (*m/z*)
Ser9	TT**S[Pho]**FAESCKPVQQPSAFGSMK	763.36	790.01
Tyr216	GEPNVS**Y[Pho]**ICSR	641.29	681.28

Table 4. 1. Peptides sequences and relative mass-to-charge ratio for LC-MS/MS analysis

Figure 4. 16 shows the extracted ion chromatograms of Ser9 (left) and Tyr216 (right) relative to GSK-3β pool 1. In the chromatogram, the intensity of the signal observed at the chosen *m/z* value is plotted in a series of mass spectra recorded as a function of retention time.

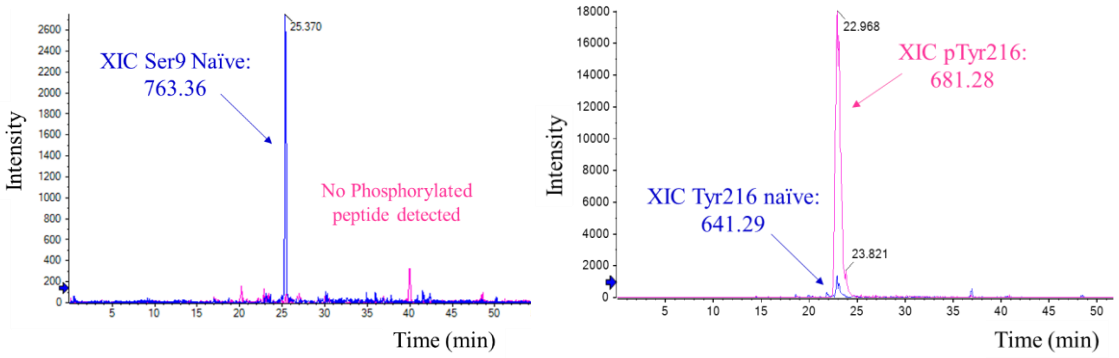


Figure 4. 16. Extracted ion chromatograms of GSK-3β pool 1.

The area of the XIC was calculated and used to express the ratio of phosphorylation. Table 4. 2 and Table 4. 3 summarize the quantification of Tyr216 and Ser9 phosphorylation calculated for each GSK-3β pool.

TTS[Pho]FAESCKPVQQPSAFGSMK			
Pool	**Naïve**	**Phos**	**%Phos**
1	32861.62	0	0%
2	970221.70	0	0%
3	1697971.01	0	0%

Table 4. 2. LC-MS/MS quantification of Ser9 phosphorylation

GEPNVSY[Pho]ICSR			
Pool	**Naïve**	**Phos**	**%Phos**
1	132273.10	2605478.53	95%
2	3919150.25	400686.65	9%
3	1901223.41	8901.34	0%

Table 4. 3. LC-MS/MS quantification of Tyr216 phosphorylation

LC-MS/MS analysis confirmed that all the considered elution pools contained GSK-3β: while in pool 1 and 2 only GSK-3β was identified, in pool 3 four other proteins from Trichoplusia ni (H5)

cells and eleven proteins from baculovirus were found besides GSK-3β. The lack of purity of pool 3 was also observed by SDS-PAGE (Figure 4. 11). The analysis also revealed a significantly different phosphorylation pattern: in pool 1, 95% of Tyr216 residues were phosphorylated, while Ser9 phosphorylation was not detected. In pool 2, only 9% of Tyr216 were phosphorylated and Ser9 was not phosphorylated. In pool 3 phosphorylation was absent both on Tyr216 and Ser9. These results correlated with the previously performed activity assays, confirming that enhanced GSK-3β activity is due to the phosphorylation on Tyr216. Surprisingly, phosphorylation of residue Ser219 was detected in pool 3. Ser219 is part of the P-2 binding site, together with Arg220 and Tyr221 (16). So far, there is no evidence of a role played by phosphorylation of Ser219 in the regulation of GSK-3β catalytic activity and very little is known about specific signaling pathways that mediate this modification.

4.3.3. Characterization of GSK-3β stability

In parallel to the characterization of GSK-3β catalytic activity and phosphorylation pattern, a study on thermal stability of the protein was also carried out with Thermal Shift Assay (TSA). TSA is a biophysical technique that measures the temperature at which a protein unfolds through the binding of a fluorescent probe to the exposed hydrophobic surfaces. In TSA, the fluorescence of the solution is monitored while the solution is heated: when the protein begins to unfold, the hydrophobic core is exposed and the signal increases until all protein molecules are denatured (81). In this way, the temperature of hydrophobic exposure at which half of the protein population is unfolded, T_m, is determined. In general, a higher melting point indicates a higher protein stability. A higher thermal stability usually implies a lower dynamisms and heterogeneity, which represents an advantage for crystallography. Moreover, a higher melting point of a protein also translates into a greater stability at lower temperatures, which can result in higher protein purification yields and an increased probability of crystallization.

	T_m (°C)
Pool 1	49.7 ± 0.6
Pool 2	41.7 ± 0.6
Pool 3	42.0 ± 0.0

Table 4. 4. Melting temperatures (T_m) of GSK-3β pools. Values are the average of three replicates.

Table 4. 4 reports melting temperatures values measured with TSA. A difference of around 8°C in melting temperature was observed between first pool (T_m= 49.7 °C) and second and third pool (T_m= 41.7 °C and 42.0 °C, respectively). These results, together with NMR studies and LC-MS/MS analysis, suggested that GSK-3β pool 1 is the most active and stable pool due to the phosphorylation of Tyr216. Therefore, this protein pool was selected for following experiments.

4.4. Validation of a novel Electrostatics Adiabatic Bias Molecular Dynamics (elABMD) method to predict protein-ligand unbinding kinetics

Despite the importance of kinetic characterization of protein-ligand interactions in the early phases of drug discovery pipeline, kinetics studies are usually left to later stages of lead optimization. While there are many experimental techniques that can directly estimate kinetic rate constants, it is challenging to optimize kinetic profiles depending on the height of the highest energy barrier (i.e. the rate-limiting step) separating the bound and unbound states. Moreover, experimental techniques able to determine (un)binding kinetics are difficult to transfer to high-throughput industrial applications.

To rationally design drug binding kinetics, it is important to get insights into the molecular determinants of drug-target interactions. In fact, while the concepts underlying rational optimization of binding affinity are relatively well understood, the same is not true for binding kinetics (48).

Molecular-dynamics (MD)-based methods can be applied to evaluate kinetic rate constants, but are limited by the timescales of dissociation times. Many different solutions have been adopted to predict unbinding kinetics, but they usually lack in providing atomic-resolution mechanistic information about dissociation pathways.

In this context, our group developed a novel method based on adiabatic bias molecular dynamics (MD) coupled with an electrostatics-like collective variable (dubbed elABMD) for sampling protein-ligands dissociation events in a fast and efficient way. The collective variable was chosen able to monitor the electrostatic potential between interacting molecules, instead of the RMSD of the bound/unbound state of the ligand, in order to extrapolate mechanistic and path information on unbinding events.

ABMD is a simulation technique that generates molecular dynamics trajectories by connecting points in conformational space separated by activation barriers (82). It was first used to simulate and study the structural changes occurring upon unfolding of lysozyme. This technique is applied to simulate conformational transformations in complex molecular systems and provides a realistic description of their evolution to an external perturbation.

The novel method was prospectively applied to a congeneric series of GSK-3β inhibitors in order to test its predictivity power. For this purpose, a recent paper by Berg and colleagues (35) was considered. In this work, a series of pyrazine derivatives was designed covering a wide chemical space, allowing the investigation of the structural basis of the potency against GSK-3β and the selectivity versus the Cyclin-Dependent Kinase 2 (CDK2). Among these compounds, a series of six strictly congeneric ATP-competitive GSK-3β inhibitors was selected (Figure 4. 17): of them, three co-crystallized complexes (PDB IDs 4ACG, 4ACD, 4ACH) whose ligands showed significantly different experimental potency (K_i) were chosen. Three additional compounds were selected from the same chemical series in order to maintain high levels of structural similarity and widening the potency range (Table 4. 5). Compounds C22, C44 and C50 were identified as valid candidates for inclusion in the series. All ligands were purposely synthetized within our group at the Italian Institute of Technology (IIT) using a synthetic strategy adapted from Berg et al (35).

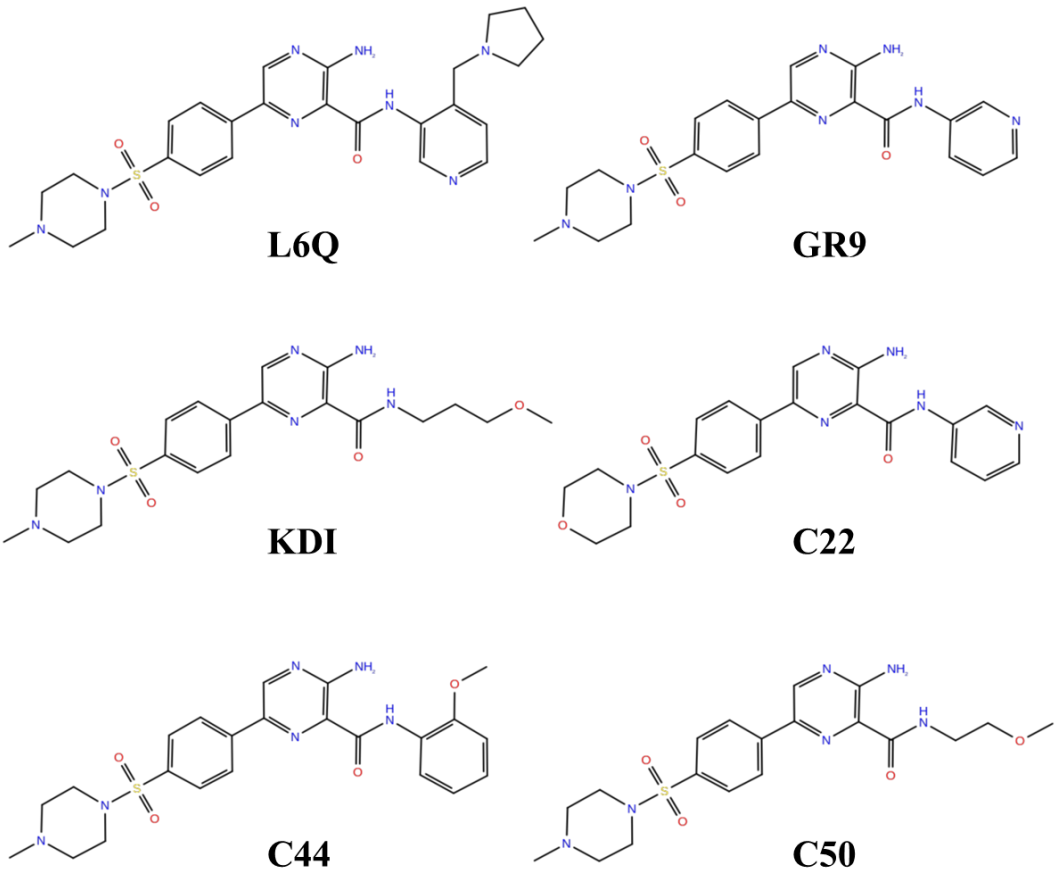


Figure 4. 17. Chemical structures of the selected GSK-3β inhibitors

Compound	K_i (nM)
L6Q	0.22
GR9	4.9
KDI	22
C22	0.67
C44	12
C50	90

Table 4. 5. Experimental potencies (Ki) of selected GSK-3β inhibitors (35)

These compounds, despite having a very similar chemical structure, display a significant difference in potency values. For example, KDI hydrophobic tail is one carbon atom longer compared to C50, and this difference cause potency to drop by a factor of 4. Morpholine group in compound C22 determines the loss of the positive charge, causing a potency increase by a factor of 7 compared to its methylpiperazine analogue GR9. These observations highlight the fact that small differences in the chemical structure significantly affect the mode of action, and, presumably, also the kinetics of binding.

In order to provide reliable starting binding poses, X-ray co-crystal structures of GSK-3β in complex with C22, C44 and C50 were also generated. The kinetic profile was experimentally characterized with surface plasmon resonance (SPR) assays to determine k_{off} values and provide a validation of compounds ranking and k_{off} values calculated with the electrostatic adiabatic molecular dynamic (elABMD) algorithm.

4.4.1. Determination of k_{off} values with SPR

Surface plasmon resonance (SPR) is a biophysical technique that exploits an optical method to measure the refractive index near a sensor surface. This surface forms the floor of a small flow cell, through which an aqueous solution passes under continuous flow. In order to detect a protein-small

molecule interaction, the protein is immobilized onto the sensor surface. The small molecule is injected in aqueous solution through the flow cell, under continuous flow. The binding of the small molecule results in an increase of the refractive index, which is measured in real time.

The result from a SPR experiment is a sensorgram, a plot of the response (in resonance units, RUs) against the time. The association phase, in which the compound binds to the target, is followed by the plateau, in which the equilibrium (steady state) is achieved. After the injection is stopped, the compound starts to dissociate (dissociation phase). The fitting of the curves corresponding to the association and dissociation phases allows the calculation of the kinetic constants K_D, k_{on} and k_{off}.

Immobilization

Immobilization of the protein onto the sensor chip surface, together with regeneration, is often the most challenging step in the optimization of an SPR experiment. It is critical to find immobilization conditions that preserve the native ligand-binding properties of the protein without compromising its structure, conformation, and binding-site accessibility. First immobilization attempts were performed with an affinity capture approach, employing the nitriloacetic acid (NTA) – nickel technique developed by Hoffmann-LaRoche. In this method, the imidazole-side chain of histidine coordinates with surface attached NTA-nickel complexes, allowing a reversible immobilization. This protocol however did not provide a stable baseline, and a gradual dissociation of the protein from the surface was observed, leading to a loss of surface activity. To improve the chip surface stability, an approach based on the capture-coupling protocol proposed by Rich et al. (83) was employed. The method is a hybrid of capture and amine coupling chemistry, where the His tag serves to preconcentrate and orient the protein onto the surface for subsequent covalent crosslinking via the activated carboxyl groups. This strategy allowed to obtain a stable baseline for subsequent binding assays. Figure 4. 18 shows a sensorgram of the final immobilization protocol. FC2 (in red) was typically used as the reference channel, while GSK-3β was immobilized on FC1 (in blue). GSK-3β was injected after $NiCl_2$ in order to orient the protein onto the surface by exploiting the bond with the his-tag, then EDC/NHS and ethanolamine allowed the covalent crosslinking with the surface.

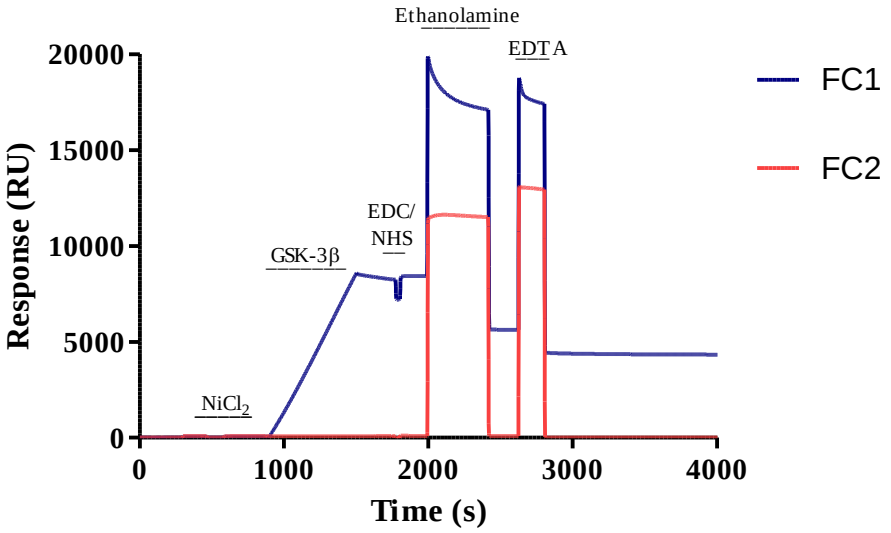


Figure 4. 18. SPR sensorgram of GSK-3β immobilization protocol

Regeneration

Proper regeneration of the surface between injections of compound is important to ensure the surface remains stable throughout the entire kinetic experiment (84). Determining an optimal surface regeneration protocol is usually a challenging step that requires a significant amount of trials. Regeneration is often achieved by the application of high- or low-pH solutions to the system. The change in pH alters the relative charges between the compound and the protein and the tertiary structure of the protein, helping the dissociation from the receptor. However, this strategy may cause irreversible damage to the protein. An alternative is to alter the ionic strength of the environment using strong electrolytes such as NaCl, which is a preferable approach when the protein is particularly sensitive (85). Since ligands L6Q, GR9 and C22 displayed a very high residence time, it was necessary to identify a suitable regeneration strategy. The first approach was to inject a solution of 10 mM glycine-HCl at different pH values (1.5-3.5) for 30 seconds, but after the injection, GSK-3β lost the ability to bind the compound. A milder solution of 1 M NaCl was injected to exploit the high ionic strength, but it was not effective in the dissociation of the strongest binding compounds. After many trials, the best regeneration conditions were obtained with the chaotrope DMSO. A solution of 15%

DMSO in H_2O preserved GSK-3β activity while completely dissociating all bound ligands, even the ones with slow dissociation rates.

Binding assay

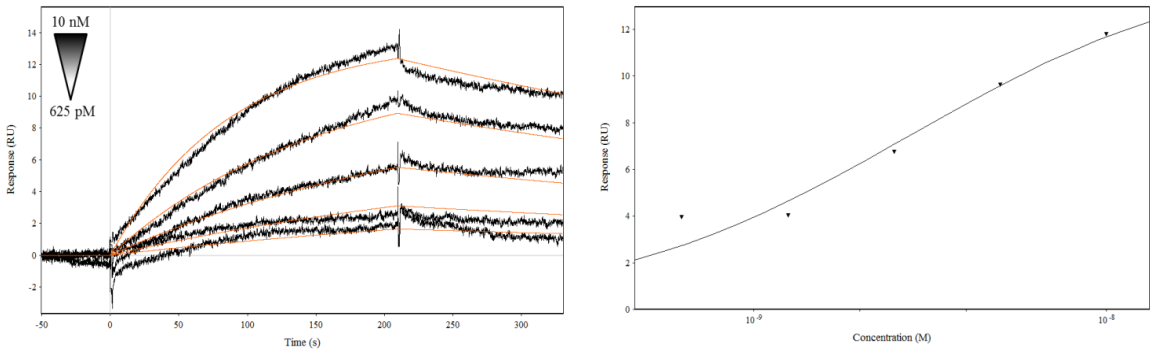
For binding assay, a buffer containing Tris-HCl was preferred to Hepes to avoid possible competition with the analytes. Tween20 was added to prevent hydrophobic based non-specific interactions and to prevent compound losses due to binding to tubing. For each compound, the top concentration was carefully optimized to obtain maximum accuracy, and from that, serial dilutions were prepared and tested. Figure 4. 19 shows SPR sensorgrams of binding assays of each compound generated during Qdat analysis. On the left, response (RU) is recorded as a function of time. The association phase is followed by the dissociation phase, allowing the calculation of kinetic dissociation constants, K_{Ds}. In the affinity plot (on the right), response (RU) is recorded as a function of compound concentration (M) allowing the extrapolation of the dissociation constant at steady state. The responses at equilibrium all fitted well to a simple 1:1 binding model. It is important to point out that high refractive index components such as DMSO alter the SPR signal causing significant shifts. For this reason, DMSO is usually added to the binding buffer at maximum concentration of 5%. The injection of DMSO as regeneration solution caused bulk shifts that were taken into account during the analysis of binding data, however they were still visible as signal peaks at the end of the association phase.

By looking at the SPR results, we observed that this series of compounds displayed a moderately wide range of association and dissociation rates. Compounds C50 and KDI displayed very fast dissociation rates, and reached equilibrium rapidly during the association phase. On the contrary, L6Q displayed very slow association and dissociation from the target. An agreement with potency values was also found. A summary of all binding constants is provided in Table 4. 6.

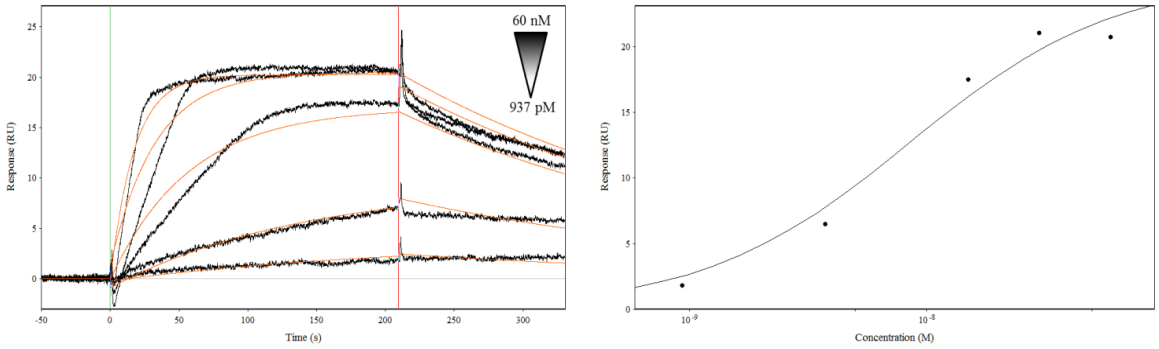
Cpd	$K_{D\ st\text{-}st}$ (nM)	$K_{D\ kin}$ (nM)	k_{on} (s^{-1} M^{-1})	k_{off} (s^{-1})	$t_{r,exp}$ (s)
L6Q	2.79 ± 0.04	1.64 ± 0.01	1.001 ± 0.003E+06	1.64 ± 0.01E-03	609.76

GR9	8.4 ± 0.1	4.29 ± 0.02	8.97 ± 0.03E+05	3.85 ± 0.02E-03	259.74
KDI	17.75 ± 0.09	10.90 ± 0.02	8.19 ± 0.01E+05	8.925 ± 0.006E-03	112.04
C22	2.81 ± 0.04	1.53 ± 0.01	1.667 ± 0.005E+06	2.55 ± 0.02E-03	392.16
C44	31.4 ± 0.6	18.93 ± 0.05	2.441 ± 0.006E+05	4.565 ± 0.005E-03	219.06
C50	45.9 ± 0.3	42.3 ± 0.2	1.286 ± 0.005E+05	5.4400 ± 0.0001E-02	18.38

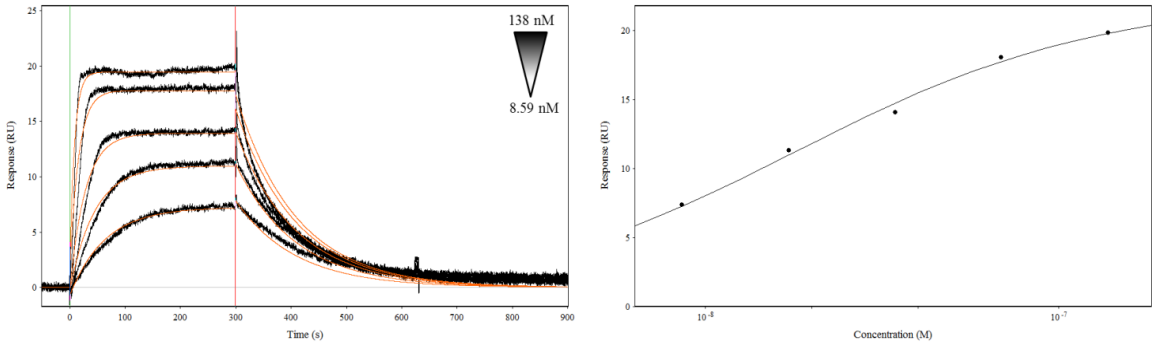
Table 4. 6. Experimental SPR kinetic data of GSK-3β ATP-competitive inhibitors. $K_{D\ st\text{-}st}$ = K_D steady-state. $K_{D\ kin}$ = K_D kinetics. $t_{r,exp} = 1/k_{off}$.



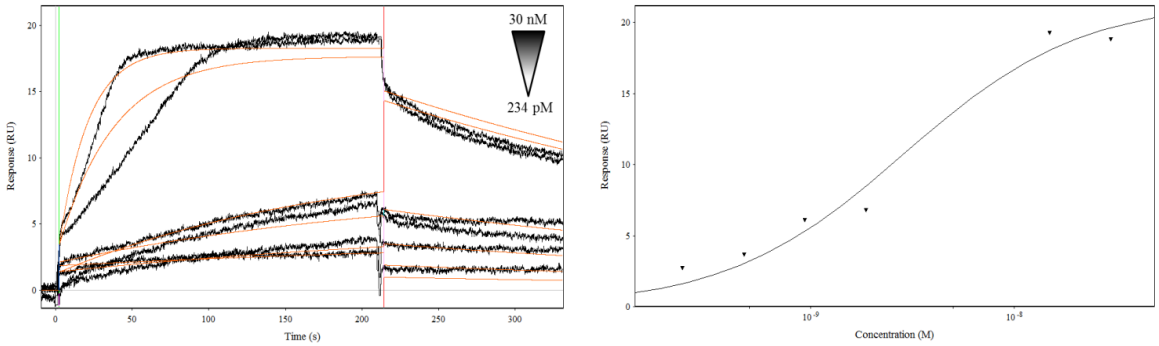
L6Q



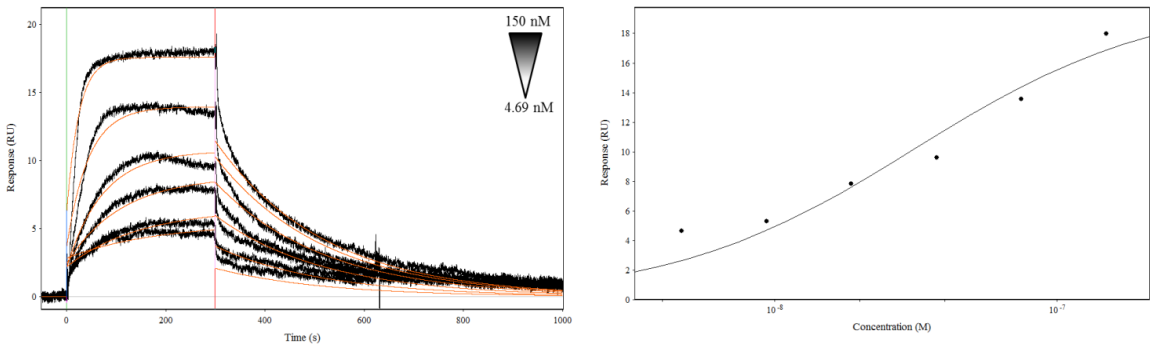
GR9



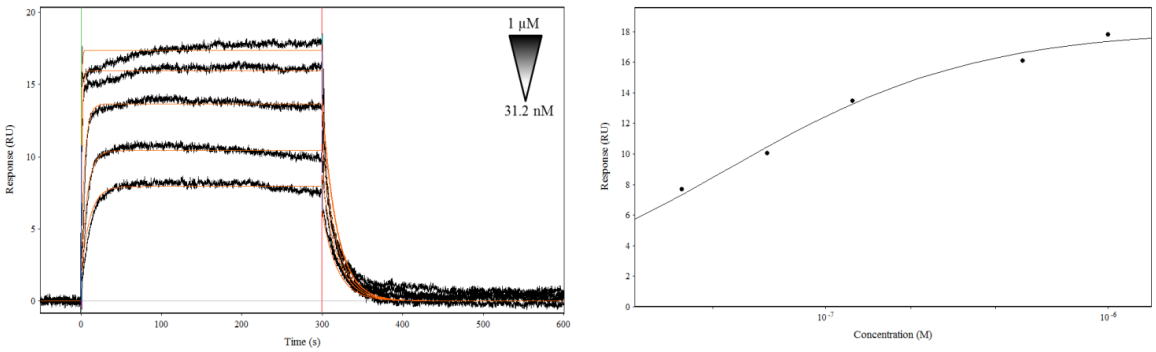
KDI



C22



C44



C50

Figure 4. 19. SPR sensorgrams of GSK-3β ATP-competitive inhibitors

4.4.2. TSA on ATP-competitive inhibitors

To further confirm the binding of compounds C44, C22 and C50 to GSK-3β, thermal shift assay (TSA) was performed. The protocol was the same used for the characterization of the different GSK-

3β populations, but in this case, the protein was incubated with 3x molar excess of the inhibitor and compared to the apo form. The assay was performed with phosphorylated (Tyr216) GSK-3β.

This assay is routinely used in drug-discovery to identify new ligands by testing the effect of the inhibitor on protein stability. In fact, if a ligand stabilizes or destabilizes the fold upon binding, there will be a change in the transition midpoint for thermal unfolding (T_m) (86). Most ligands stabilize proteins upon binding, causing an increase in the protein melting temperature. Usually, equilibrium binding ligands increase protein thermal stability by an amount proportional to the concentration and affinity of the ligand (87).

Figure 4. 20 shows the fluorescence profile of GSK-3β as a function of temperature of the apo protein and upon incubation with compounds C44, C22 and C50. T_m values are reported in Table 4. 7.

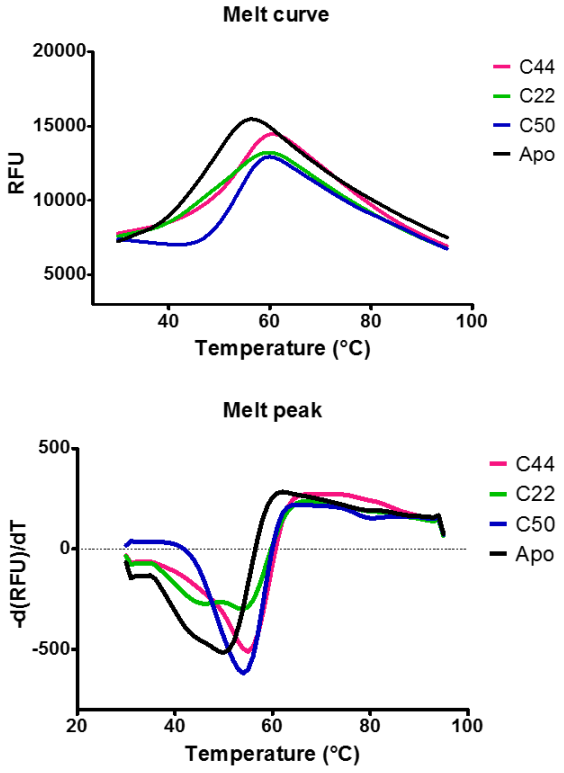


Figure 4. 20. TSA melting curves and melting peaks of GSK-3β apo and after incubation with ATP-competitive inhibitors C44, C22 and C50. Curves are the average of three replicates.

	T_m (°C)
Apo	49.7 ± 0.6
C44	55.0 ± 0.0
C22	54.0 ± 1.4
C50	54.0 ± 0.0

Table 4. 7. Melting temperatures (T_m) of GSK-3β apo and after incubation with C44, C22 and C50. Values are the average of three replicates.

As shown from the fluorescence profile and from the variation of the melting temperature, all ATP-competitive inhibitors had an impact on GSK-3β stability: C22 and C50 caused an increase of 4°C in GSK-3β T_m compared to the apo protein, while addition of C44 resulted in an increase of 5°C. These results well correlated with their affinities in the nanomolar range (35), and further confirmed the binding already observed with SPR experiments.

4.4.3. Crystallization of ATP-competitive inhibitors C22, C44 and C50

For three out of the six ligands (L6Q, GR9 and KDI), a crystallographic structure was already available (35). We decided therefore to pursue the structural study of the remaining compounds (C22, C44 and C50), as a reliable configuration is mandatory to run unbinding simulations with MD-related approaches.

In order to obtain well diffracting crystals, initial trials were performed with Hampton Research PEG/Ion Screen I and II, dispensing 0.1 µL protein and 0.1 µL reservoir using the nanoliter liquid handling system Mosquito® TTP (Labtech). Different crystallization conditions were tested in parallel with both phosphorylated (pTyr216) and unphosphorylated full-length his-tagged GSK-3β (pool 1 and pool 2, respectively), in the apo form and in complex with ATP-competitive inhibitors. Protein was concentrated to a maximum of 11.1 mg/mL. Co-crystallization was the first method of choice, as it is usually used when the compounds have low solubility. Sitting drop and hanging drop techniques were taken into considerations, however hanging drop technique did not produce any

crystal. In sitting drop plates, numerous crystals were observed within 4 days of incubation at 20°C under different crystallization conditions.

In general, most crystals appeared under neutral pH conditions (pH 6.5 – 7.5). Phosphorylated GSK-3β (pTyr216) produced a significantly higher number of crystals compared to the unphosphorylated form. This is coherent with TSA results described in paragraph 4.3.3, which highlighted a higher stability of the phosphorylated population.

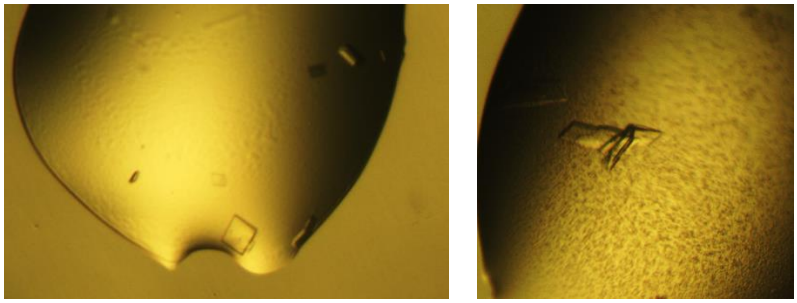


Figure 4. 21. Examples of phosphorylated (pTyr216) apo GSK-3β crystals. Left: (0.02 M citric acid, 0.08 M BIS-TRIS propane / pH 8.8), 16% PEG 3350. Right: 0.1 M sodium formate pH 7.0, 12% PEG 3350

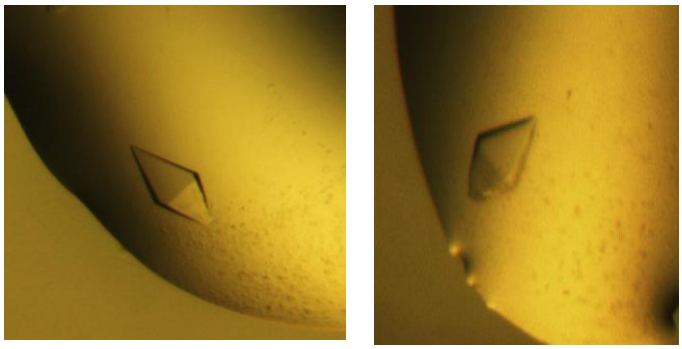


Figure 4. 22. Examples of unphosphorylated apo GSK-3β crystals. Left: 0.1 M Hepes pH 7.5, 2% Tacsimate pH 7.0, 20% PEG 3350. Right: 0.1 M Tris pH 8.5, 2% Tacsimate pH 8.0, 16% PEG 3350.

Differences in the shapes were also observed, as phosphorylated GSK-3β crystals appeared smaller and flat-shaped, or two-dimensional plates overlaying each other (Figure 4. 21). Contrariwise, unphosphorylated crystals were either needle- or diamond-shaped (Figure 4. 22). The optimization of crystallization conditions was then focused on the phosphorylated form, by varying drops volumes, PEG concentration and pH of the buffer.

Crystal structure of C22 bound to GSK-3β was finally obtained by sitting drop co-crystallization in 8% tacsimate pH 7.0, 20% PEG 3350 with a resolution of 2.3Å. Crystal structure of C44 was obtained by sitting drop co-crystallization in 0.1M Bis-Tris pH 6.5, 2% tacsimate pH 6.0, 20% PEG 3350 with a resolution of 2.5Å. Crystallization of C50 turned out to be more challenging than expected. For this compound, sitting drop soaking technique was employed, and crystal structure was obtained in 0.2M potassium fluoride, 22% PEG 3350 at 3.2Å resolution. Images of diffracting crystals of ATP-competitive inhibitors are depicted in Figure 4. 23, and a summary of final crystallization conditions is reported in Table 4. 8.

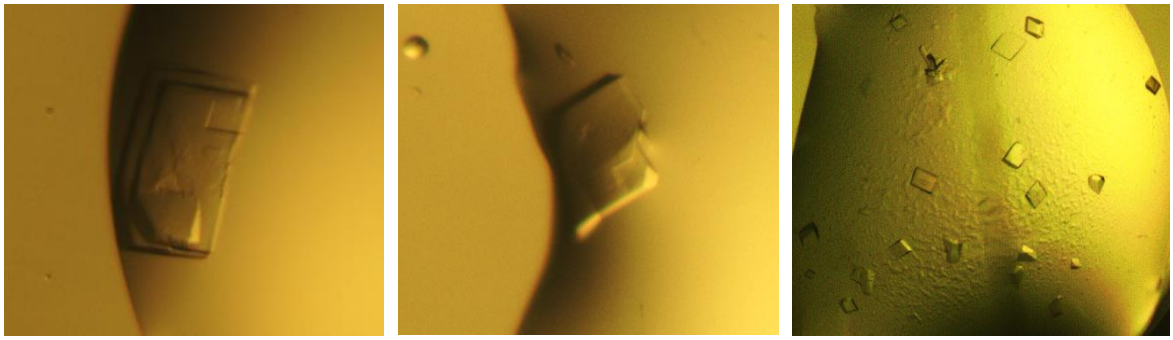


*Figure 4. 23. Optical microscope images of crystals grew from 8% tacsimate pH 7.0, 20% PEG (**ligand C22, left**); 0.1M Bis-Tris pH 6.5, 2% tacsimate pH 6.0, 20% PEG (**ligand C44, center**); 0.2M potassium fluoride, 22% PEG (**ligand C50, right**)*

	C22	C44	C50
Protein sequence	Full length (1-420), N-terminal His_6 tag, Tyr216 phosphorylated	Full length (1-420), N-terminal His_6 tag, Tyr216 phosphorylated	Full length (1-420), N-terminal His_6 tag, Tyr216 phosphorylated
Crystallization technique	Sitting drop co-crystallization	Sitting drop co-crystallization	Sitting drop soaking
Protein buffer	20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT	20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT	20 mM HEPES pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT
Crystallization condition	8% tacsimate pH 7.0 20% PEG 3350	0.1 M Bis-Tris pH 6.5 2% tacsimate pH 6.0 20% PEG 3350	0.2 M potassium fluoride 22% PEG 3350
Protein conc.	10 mg/mL	10 mg/mL	7.5 mg/mL
Inhibitor conc	3X molar excess	3X molar excess	3X molar excess
Drops	0.1µL protein+0.1µl reservoir	0.5µL protein+0.5µL reservoir	0.5µL protein+0.5µL reservoir

Table 4. 8. Final crystallization conditions for ATP-competitive inhibitors

Data acquisition details and refinement procedure for each crystal structure is described in Materials and methods section. Diffraction data corresponding to complex structures of ligand C22 and C44 were relatively better compared to C50 (Table 4. 9). Molecular replacement placed two protein molecules in the asymmetric unit in P1 space group for structures corresponding to ligand C22 and C44. In case of ligand C50, a single molecule was placed in the asymmetric unit in C2221 space group. Although each crystal varied in terms of quality of data, unambiguous electron density was observed corresponding to each ligand along with the residues surrounding the ligand. Coordinates were deposited in the PDB with IDs 6HK3, 6HK4, 6HK7.

All structures confirmed the binding mode of the compounds in the ATP pocket of GSK-3β. The X-ray crystal structure of compound C44 bound to GSK-3β (Figure 4. 24) indicated that nitrogen N3 of the pyrazine ring of C44 establishes a hydrogen bond with the backbone nitrogen of Val135 and

the main chain oxygen of Asp133 in the ATP site. Furthermore, amine N6 is hydrogen bonded to the peptide oxygen of Asp133. The amide oxygen O3 of C44 forms a water-mediated link to the peptide nitrogen of Asp200. Pyperazine N1 is showing water-mediated hydrogen bond to OH group of Tyr134.

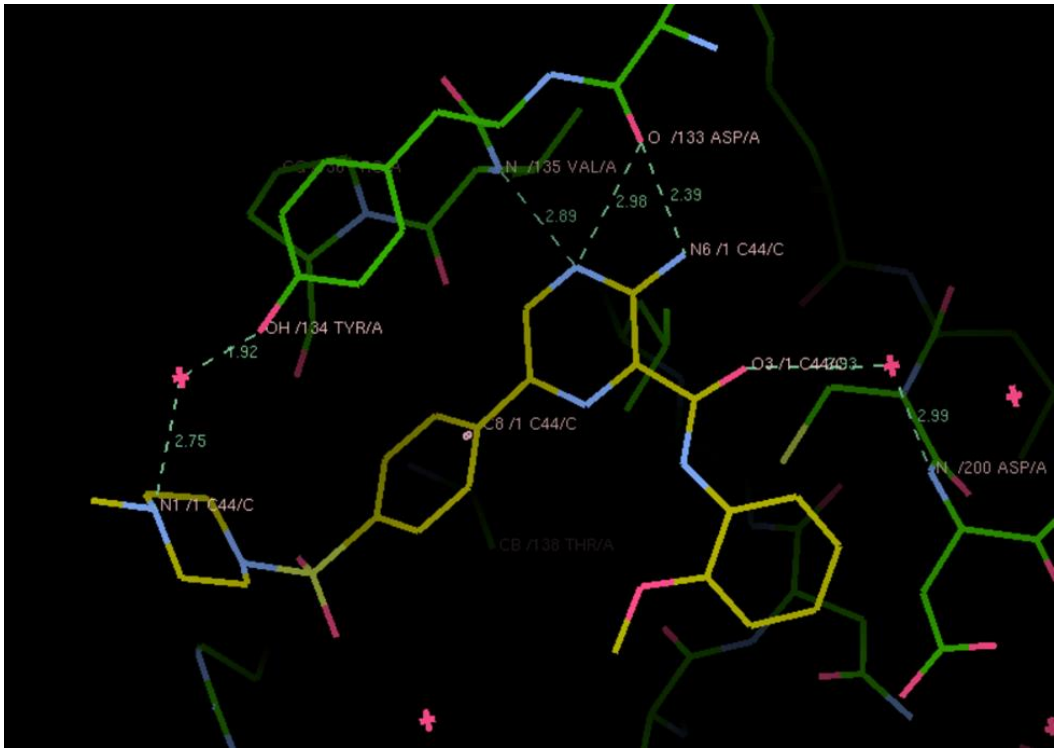


Figure 4. 24. Crystallographic pose of ligand C44 bound to GSK-3β (PDB ID 6HK3).

By looking at the crystallographic pose of C22, it can be observed that most of hydrogen bonding interactions of C22 are similar to the ones of C44 (Figure 4. 25). C22 establishes an extra hydrogen bond of sulphonyl oxygen O3 with side chain nitrogen NH1 of Arg141.

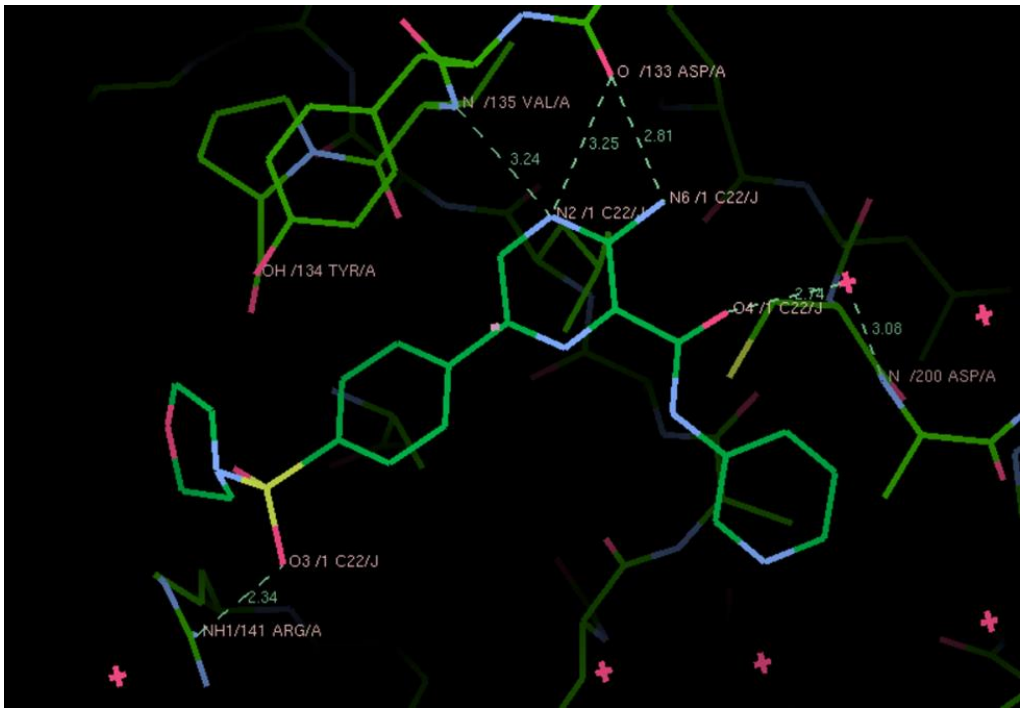


Figure 4. 25. Crystallographic pose of ligand C22 bound to GSK-3β (PDB ID 6HK4).

Compared to C22, compound C50 differs in a terminal pyridine ring that is replaced by a methoxypropyl group. However, despite this structural difference, the overall binding mode of C50 is very similar to the pose of C22 and C44 (Figure 4. 26). The same orientation was also observed in X-ray crystal structures of pyrazine compounds GR9 (PDB ID 4ACD) and L6Q (PDB ID 4ACG). L6Q has the polar pyrrolidine ring extending in the ribose binding site, while in the crystal structure of compound GR9 this site is occupied by a water molecule, resulting in a 22-fold decrease in potency compared to L6Q. Moreover, although the pyridine nitrogen is in the same position, the piperazine ring is shifted slightly up and out toward the solvent area for compound L6Q.

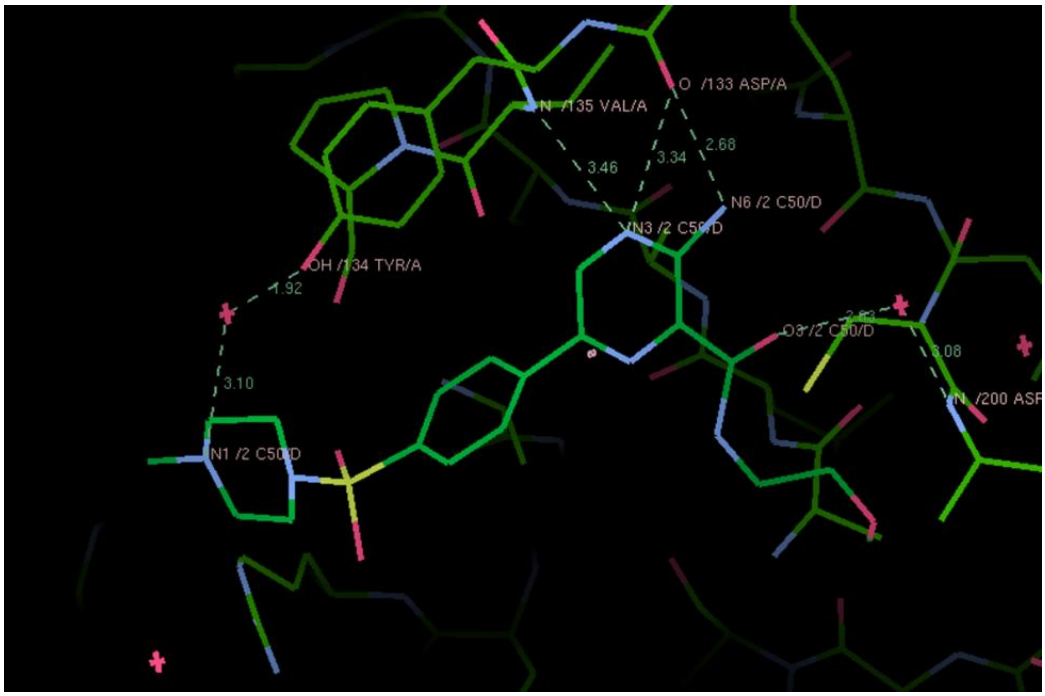


Figure 4. 26. Crystallographic pose of ligand C50 bound to GSK-3β (PDB ID 6HK7).

Overlapped binding poses of C22, C44 and C50 are depicted in Figure 4. 27.

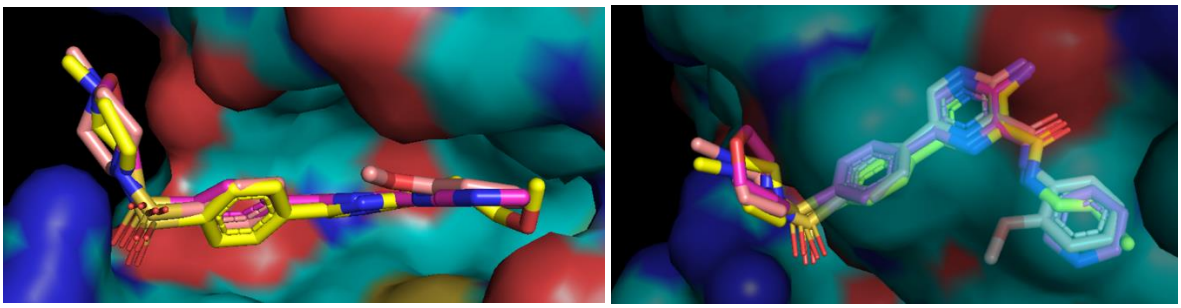


Figure 4. 27. Side view (left) and top view (right) of overlapped binding poses of compounds C22, C44 and C50

Details of crystallographic data processing and analysis are summarized in Table 4. 9.

Data collection			
	6HK3 (Ligand C44)	**6HK4 (Ligand C22)**	**6HK7 (Ligand C50)**
Space group	P1	P1	$C222_1$

Cell dimensions a,b,c (Å)	64.32 67.12 67.24	63.19 67.32 67.2	85.979 103.727 108.92
Wavelength (Å)	1.0	1.0	1.0
Resolution (Å)	42.98 - 2.34 (2.34-2.42)	46.25 - 2.2 (2.2-2.7)	51.86 - 3.0 (3.0 - 3.18)
Unique reflections	47982 (2048)	50974 (3771)	21008 (3208)
†R_{merge}	0.1 (1.454)	0.1 (1.536)	0.255 (0.986)
<I/σ>	5.2 (0.5)	5.3 (0.5)	4.9 (1.3)
Completeness (%)	97.1 (97.1)	95.2 (81.1)	99.5 (98.8)
Redundancy	2.4 (2.5)	2.0 (1.8)	5.4 (5.7)
$CC_{1/2}$	0.99 (0.34)	0.99 (0.268)	0.97 (0.75)
Wilson plot B-factor ($Å^{-2}$)	44.5	46.436	50.282
Refinement			
Resolution (Å)	42.977 - 2.35	43.67 - 2.5	46.87 – 3.2
No. of reflections	43734	35615	7856
ǂRwork / +Rfree	0.2041/.2335	0.2098/0.2498	0.200/0.285
Total number of atoms	5967	6023	2659
Protein atoms	5638	5642	2615
Water	213	245	4
Glycerol	24	36	6
DMSO	8	24	4
Malonic acid	14	14	-
Inhibitor	68 (C44)	62 (C22)	30 (C50)
Chloride Cl^-	2	-	-
RMSD bonds (Å)	0.009	0.009	0.009
RMSD angles (°)	0.981	1.075	1.717

†R_{merge} = $\Sigma_{hkl}\Sigma_i|I(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ *is the intensity of the* i^{th} *measurement and* $\langle I(hkl)\rangle$ *is the mean intensity for that reflection.*

‡R_{work} = $\Sigma_{hkl}||F_{obs}| - |Fc_{alc}||/\Sigma_{hkl}|F_{obs}|$, where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure-factor amplitudes, respectively.

+$\mathbf{R_{free}}$ was calculated with 5.0% of reflections in the test set.

Table 4. 9. Details of X-ray crystallographic data collection and refinement statistics

4.4.4. Computational prioritization and comparison with experimental SPR data

After establishing a reliable starting configuration for compounds C22, C44 and C50, computational simulations were run on the selected pyrazine GSK-3β inhibitors with novel elABMD algorithm. Simulation accuracy was set high in order to minimize the system perturbation, decreasing the constant of force. An extensive statistics of 20 independent replicas of 30 ns for each system was collected to increase the robustness of the predictions, with a total of 4.2 μs of production runs. In Table 4. 10 are reported the unbinding-time-based ranking correlations resulted from elABMD mean unbinding times, in comparison to experimental kinetic data collected with SPR.

	elABMD		**Experimental**	
Cpd	$\mathbf{t_{r,\ mean}}$	**Rank** $\mathbf{t_{r,\ mean}}$	$\mathbf{t_{r,\ exp}}$	**Rank** $\mathbf{t_{r,\ exp}}$
L6Q	28.64 ± 0.60	1	609.76	1
GR9	26.53 ± 1.56	3	259.74	3
KDI	25.72 ± 1.40	4	112.04	5
C22	28.45 ± 0.77	2	392.16	2
C44	24.94 ± 1.70	5	219.06	4

C50	18.75 ± 2.02	6	18.38	6

Table 4. 10. Mean estimation of elABMD unbinding times ($t_{r,\ mean}$) and experimental kinetic data

elABMD unbinding times ($t_{r,\ mean}$) are expressed in ns and reported with an estimation of the error computed with a bootstrapped procedure. Experimental residence times are expressed in s.
elABMD were in good agreement with experimental residence times: three groups of GSK-3β inhibitors with increasing predicted unbinding times were defined. The first group included the most potent inhibitors of the series (L6Q, GR9 and C22), the second group included compound KDI and C44, and the third group included the weakest compound C50.

After establishing that predicted unbinding-time-based ranking correlations and experimental data were in good qualitative agreement, an analysis on how subtle structural modifications affect the unbinding mechanism was performed to provide initial structure-kinetics relationships. By observing the collected trajectories, it was noticed that all inhibitors explored a homogeneous unbinding path directed toward the small N-lobe, presumably preceded by a slight rotation of the scaffold in the ATP-binding pocket. Binding poses were initially destabilized by the upward movements of the methylpiperazine group or by the fluctuations of substituents to the amide group leading to structural rearrangements of the glycine-rich loop. It was also observed that the calculated unbinding time increased when the motions of the functional groups connected to the amide group were limited by the establishment of hydrogen bonds with Lys85, or by intramolecular interactions with the adjacent amide nitrogen atom. In compounds L6Q, GR9 and C22 the nitrogen atom of the pyridine ring interacted directly with the Lys85 sidechain (Figure 4. 28). This happened for the terminal methoxy group of C50 to a lesser extent.

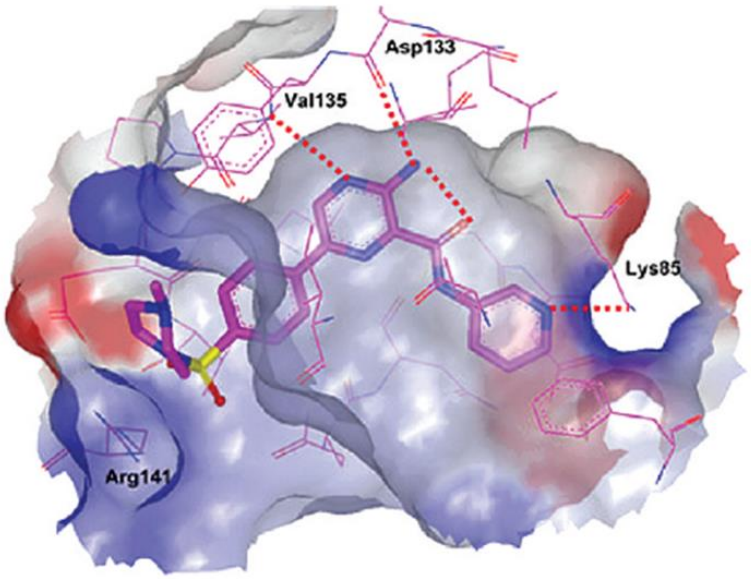


Figure 4. 28. X-ray crystal structure of compound GR9 in the ATP binding site of GSK-3β (35)

In the bound state, KDI and C44 were stabilized by intramolecular interactions. In both cases, the resulting dissociation mechanisms were dependent on structural rearrangements of the glycine-rich loop opening the flexible binding site.

To evaluate the influence of the hydrogen bond with Lys85 sidechain on unbinding kinetics, the dissociation mechanisms explored by GR9 and C44 were compared. In C44, the stability of the bound state was affected by the degree of freedom of the rotatable bond of the methoxy substituent in ortho position, facilitating the rupture of the internal hydrogen bonds. Contrariwise, GR9 maintained the hydrogen bond with Lys85. Therefore, the unbinding time was increased by the structural rearrangements of the scaffold inside the binding site, limited by hydrogen bonds with conserved Lys85. In the crystal structure of GSK-3β complexed with GR9, the ribose binding site is occupied by a water molecule. This water molecule was absent in complexes of L6Q and KDI, in which the ribose binding site was partially filled by the folded conformation of substituents to the amide group. This difference suggested a gain in potency due to the displacement of a high-energy water molecule. During the dynamics, the pyrazine ring of L6Q interacted stably with Lys85, and the pyrrolidine ring established water-mediated interactions with the C-lobe residues or with the glycine-rich loop. In contrast, the folded conformation of the methoxypropyl chain of KDI adopted an extended

conformation. This resulted in the solvent-exposure of the hydrophilic methoxy group, which allowed water molecules to partially fill the ribose binding site. This cancelled the gain in potency acquired by the displacement of the high-energy water molecule necessary for binding.

These observations deepened the structural basis of different experimental potency values, thus potentially contributing to the improvement of protein-ligand complex lifetime. It was observed that hydrophilic substituents in specific positions increased the duration of the protein-ligand complex, by reducing structural rearrangements of protein residues around the binding site and limiting the degree of accessibility of water molecules. The agreement between computational predictions and experimental measures suggests that elABMD is an efficient and innovative tool for calculating residence times in drug discovery programs, in particular in hit-to-lead and lead optimization phase, offering a qualitative description of the unbinding mechanism in a very modest computational time.

4.5. Characterization of GSK-3β allosteric inhibitors 20g and 20f

So far, only a few chemically different non-ATP competitive GSK-3β inhibitors have been reported. Most known allosteric inhibitors (e.g., manzamine A and palinurin) have a very complex chemical structure, which limits further studies, and are not commercially available. This might be one the reasons why to date, no crystal structure of GSK-3β bound to an allosteric inhibitor has yet been described, meaning that a deep structural characterization of allosteric modulation mechanism is still lacking.

Through a literature research, we identified two promising allosteric inhibitors recently described by Gao et al.(46). In this work, researchers synthetized a series of novel substituted benzothiazinones (BTOs) derivatives as GSK-3β allosteric inhibitors, designed by merging two allosteric GSK-3β inhibitors: the basic core of BTO-jb and the long side chain of VP0.7 (36). The length and flexibility of the side chain attached to the carbonyl group appeared to be crucial for inhibitory activity. Among this series, compounds 20f and 20g (Figure 4. 29) showed good potency (11.7 μM and 22.4 μM, respectively) and a moderate antiproliferative effect on ovarian cancer cell lines. In particular, compound 20g induced apoptosis and arrested the cell cycle at the G1 phase. 20g was also screened against a panel of serine/threonine kinases closely related to GSK-3β, showing lower or no inhibitory effect compared with GSK-3β.

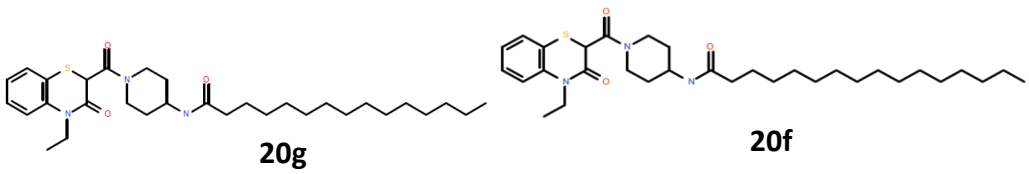


Figure 4. 29. Chemical structures of allosteric compounds 20g and 20f

In order to gain more structural details on the binding mechanisms of 20g and 20f, these compounds were chemically synthetized within our group at IIT and their allosteric modulation of GSK-3β was investigated by biophysical and structural methods.

4.5.1. Thermal Shift Assay

Thermal shift assay was carried out as described for ATP-competitive inhibitors in the previous section. However in this case, since dealing with allosteric inhibitors, GSK-3β was first incubated with AMP-PNP (a nonhydrolyzable ATP analog) for 30 minutes in order to prevent the binding of the compounds to the ATP binding site.

Results of TSA are reported in Table 4.11. Differently from ATP-competitive inhibitors, the observed shift in melting temperature compared to the apo protein was very subtle. An increase of only 1°C in the melting temperature was detected after incubation of GSK-3β with 20g and 20f compared to the T_m of the protein in the absence of the compounds.

	T_m (°C)
Apo	49.7 ± 0.6
20f	51.0 ± 0.0
20g	50.7 ± 0.5

Table 4. 11. TSA melting curves and melting peaks of GSK-3β apo and after incubation with allosteric inhibitors 20g and 20f. Curves are the average of three replicates.

4.5.2. Microscale Thermophoresis

Microscale Thermophoresis (MST) is an immobilization-free biophysical technique based on the measurement of the motion of molecules in a localized temperature gradient created (an effect called *thermophoresis*) by a highly focused infrared laser. It is a powerful tool to quantify biomolecular interactions and it offers the advantages of a fast determination of binding constants using very small amounts of samples (81). The MST protocol is described in the Materials and methods section.

Prior to the binding affinity analysis, binding check tests on 20g and 20f were carried out. Binding check experiment rapidly tests the signal to noise ratio by comparing MST signals of the target alone and the target in the presence of the ligand at a fixed concentration. The response amplitude for 20g was 14.2 and the signal to noise ratio was 13.9 (Figure 4. 30), confirming the presence of binding. However, aggregation was detected in the capillaries containing both the target and the ligand. This issue was also observed when testing 20f, for which signal to noise ratio could not be calculated.

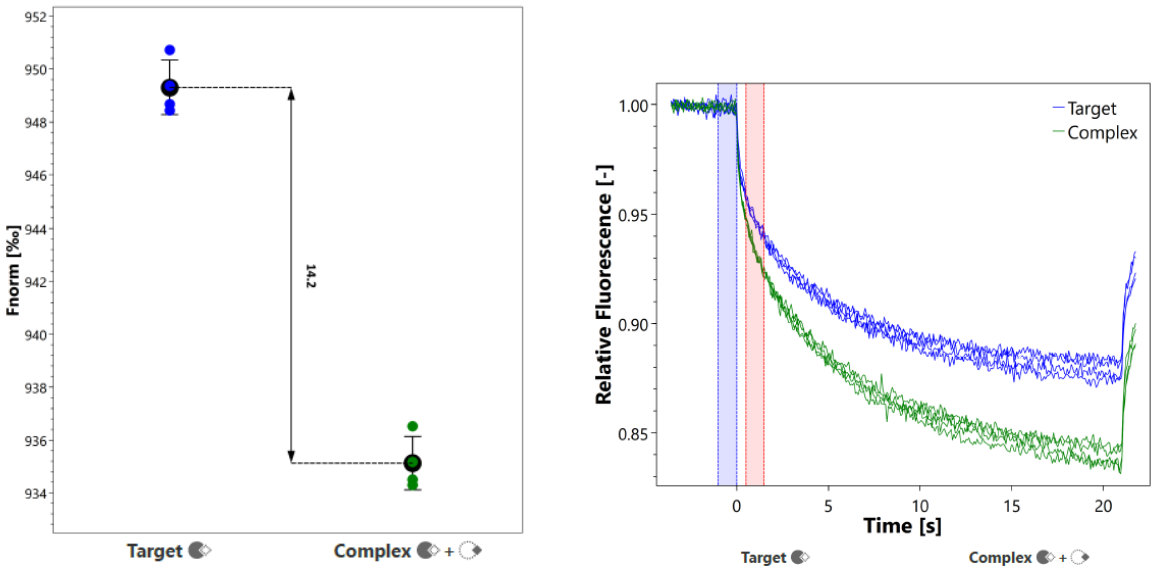


Figure 4. 30.Signal to noise ratio and MST traces of compound 20g

After the detection of aggregation, the solubility of 20g and 20f in aqueous buffer were tested with NMR (data not shown) and aggregation was confirmed even at very low µM concentration. Given the moderate potency of these allosteric compounds, further binding affinity studies could not be performed due to their very low solubility. In cell-based experiments performed by Gao at al (46), the solubility issue was probably overcome by the presence of carrier proteins or other factors present in cells and in cells growth media.

4.5.3. Crystallization trials

Crystallization trials for 20g and 20f in complex with GSK-3β were pursued with sitting drop technique by initially preparing the crystals of the protein without potential ligands. GSK-3β in the presence of a saturating concentration of AMP-PNP was mixed with mother liquor in the plate wells and incubated at 20°C until crystals appeared. Given the very low solubility demonstrated in aqueous buffer, allosteric inhibitors were directly added (from a 20 mM stock solution in 100% DMSO) on the drop containing the GSK-3β crystals and soaked for a minimum of 4 hours prior to freezing.

Crystallization plates were prepared starting from the conditions that performed better with ATP-competitive inhibitors: in particular, we focused on 0.2 M potassium fluoride, 0.2 M potassium formate, 0.2 M sodium malonate pH 7.0 and 0.1 M Hepes pH 7.5, 2% tacsimate pH 7. PEG 3350 concentration was varied from 16 to 26%. Different forms of GSK-3β were taken into consideration: phosphorylated (pTyr216), unphosphorylated and unphosphorylated without the his-tag. In this way, we also investigated whether the absence of the tag could favor the formation of protein crystals.

As observed for ATP-competitive inhibitors, GSK-3β crystals appeared under all conditions tested. However, there were differences in the shape and number of protein crystals. GSK-3β pool without the his tag displayed significantly bigger crystals with tridimensional shape (Figure 4. 31), suggesting that the presence of the tag probably hinders crystals formation. The presence of AMP-PNP, contrariwise, did not seem to affect crystallization.

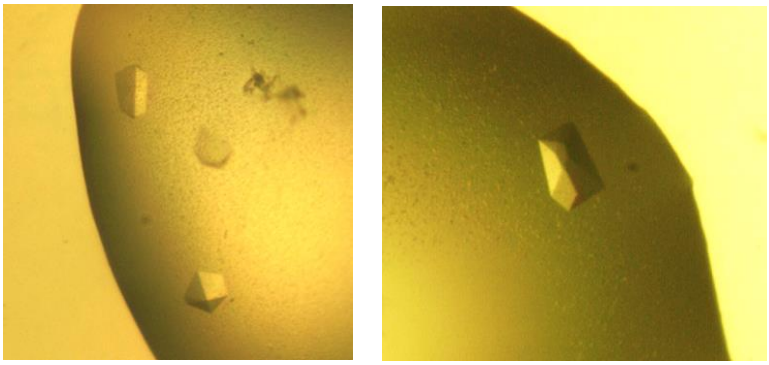


Figure 4. 31. Examples of unphosphorylated untagged AMP-PNP-GSK-3β crystals. Left: 0.2 M potassium formate, 20% PEG3350. Right:0.2 M sodium malonate, 22% PEG3350

After the soaking of the crystals with allosteric inhibitors 20g and 20f, crystals were cryo-protected in mother liquor with 20-25% glycerol and a molar excess of inhibitor, and flash-frozen in

liquid nitrogen prior to data collection. X-ray data were collected at beamline XRD1 of Elettra Synchrotron in Trieste. Unfortunately, despite the good quality of the diffraction patterns, electron densities corresponding to the inhibitors were not detected on GSK-3β surface.

4.6. Identification of novel GSK-3β allosteric inhibitors

4.6.1. Identification of allosteric pockets

Inspired by the work of Palomo and colleagues (36), we focused the research on the discovery of novel druggable binding sites on GSK-3β surface able to be modulated by small molecule interactions.

In that study, twenty-five different GSK-3-ligand complexes were considered for the calculation of cavities for ligand binding sites with software *fpocket* and *hpocket*. In the majority of the structures, the program identified between 10 and 12 different cavities, and seven pockets were highly reproducible and appeared in most of the structures analyzed. Cavities identified as cavity 1,2,3 were the only well-known binding sites on GSK-3 and corresponded to ATP-, substrate-, and axin/fratide binding sites, respectively.

In the present research, the analysis of the overall surface and communication among putative pockets was performed in collaboration with our computational group at IIT with Pocketron (74), a tool implemented in BiKi Life Science software (72). Most of currently available methods to detect protein binding pockets in static structures are limited to analyzing a priori defined pockets of interest only. This novel algorithm, differently from previous methods, allows the detection of the formation and spatiotemporal evolution of all protein pockets.

Prior to testing the algorithm, a search of tridimensional structures of GSK-3β from the Protein Data Bank (PDB) was carried out. Fifty different GSK-3β structures obtained by X-ray diffraction with a resolution better than 3Å were available at the moment of the study. Of these structures, only two were in the apo form without any small molecule or peptide bound, and one was bound to an ADP molecule. PDB structure 1J1C was chosen for the calculation as it had the most complete electron density.

A molecular dynamics simulation of 1 μs was performed on the clustering calculated by BiKi on structure 1J1C. The simulation detected over 40 pockets on GSK-3β.surface, which were mapped on 1J1C structure in Figure 4. 32.

Figure 4. 32. Pockets on GSK-3β identified with Pocketron analysis mapped on 1J1C structure

Among the most persistent ones, the algorithm identified pocket ID p6 and p1. These pockets have an average volume of 415 $Å^3$ and 96 $Å^3$, respectively. The analysis revealed a coherence with the pockets identified in the study by Palomo et al. (36), as summarized in Table 4. 12, confirming the presence of many binding sites on GSK-3β surface. The largest pocket p6 well correlates with cavity 1, which corresponds to the orthosteric site. From the Pocketron analysis it is evident that the cavity identified by Palomo as cavity 5 is in communication with the orthosteric site. This cavity is composed by pockets p1 and p10 (Figure 4. 33), which presumably contribute to the putative allosteric pocket. It has been predicted that this is also the binding site of the compound of natural origin palinurin, which is known to inhibit GSK-3β. For these reasons, this cavity has been chosen for virtual screening analysis as described in the following paragraph.

***fpocket* and *hpocket* cavities**	**Amino acids involved in the cavity**	**Pocketron cavities**
1	Lys85, Glu97, Asp133, Tyr 134, Val135, Thr138, Asn186, Leu188, Cys199, Asp200	p6
2	Phe67, Gln89, Lys94, Phe93, Asn95, Arg96, Glu97, Arg180, Lys205,	p3, p19 and p46
3	Ser215, Arg220, Arg223, Phe229	p26
4	Tyr140, Arg144, Arg148, Gln185, Ser219, Arg220, Tyr221, Tyr222, Glu249	p29 and p50
5	Met26, Thr38, Tyr56, Tyr71, Lys86, Ser119	p1 and p10
6	Glu80, Arg111, Arg113, Asp133, Val135, Asp190, Lys197	p8
7	His173, Cys178, Leu207, Arg209, Glu211, Thr235, Thr330, Tyr234, Ser236, Ser369	p7 and p2

Table 4. 12. Amino acids involved in each cavity found on GSK-3 with fpocket and hpocket and corresponding pocket IDs found with Pocketron.

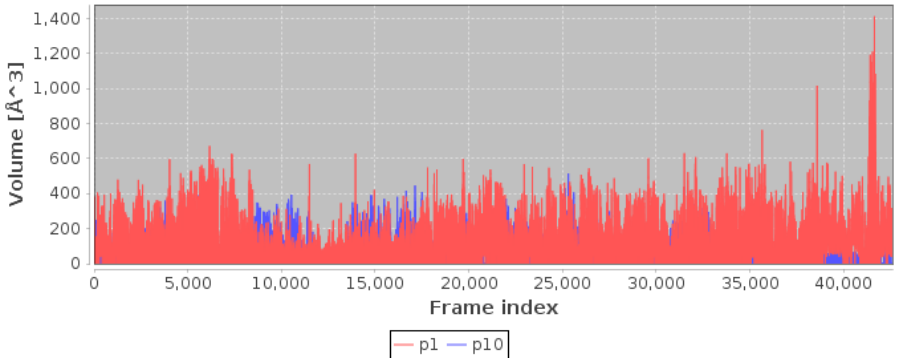


Figure 4. 33. Graph describing p1 and p10 pockets average volume (Å^3) at each time of Pocketron simulation.

4.6.2. Virtual screening

Considering the agreement between our computational study on allosteric pockets and the work by Palomo and coworkers (36), a virtual screening analysis on our in-house chemical library was performed targeting cavity 5 (Met26, Thr38, Tyr56, Tyr71, Lys86, Ser119), which was considered the most interesting pocket. This cavity corresponded to pockets p1 and p10 identified by the MD simulations in the Pocketron study previously described. In a study from 2013 (42), this cavity was also identified as a binding site for allosteric inhibitor palinurin by docking and molecular dynamics simulations. In this pocket, Lys86 forms a salt-bridge and a hydrogen bond with the tetronic ring of palinurin, and Tyr56 is hydrogen bonded to the carbonyl group of the ring. According to this study, the binding of palinurin to cavity 5 modulates the accessibility of the ATP γ-phosphate by constraining the conformation of the glycine-rich loop. These evidences point out a potential role of cavity 5 as allosteric binding site on GSK-3β surface, and therefore it was chosen for virtual screening analysis.

Virtual screening is a powerful computational technique applied in drug discovery to search libraries of small molecules in order to identify the structures that are most likely to bind to a drug target. In general, it represents a complementary alternative to the widely used high-throughput biochemical compounds screening (HTS). However, HTS is not an appropriate tool to find allosteric compounds, because the affinity of binding to the ATP outside area is usually weaker than in ATP-binding site, and the activities of this type of inhibitors are generally at micromolar level, therefore potential hits are easily abandoned off during screening (44). Virtual screening can be considered as the computational counterpart of high-throughput screening, and present some advantages in hits findings, such as velocity and cost-effectiveness.

Virtual screening was performed with Schrödinger Glide software, by centering the grid on residue Ser236. Reference crystal structures were chains A of structures 4NM5, 3IB4 and chain B of 1J1C. The procedure is described in detail in the Materials and methods section.

Compounds were selected by visual inspection and clustered based on chemical similarity. Chemical structures of the 9 virtual screening hits are reported in Figure 4. 34.

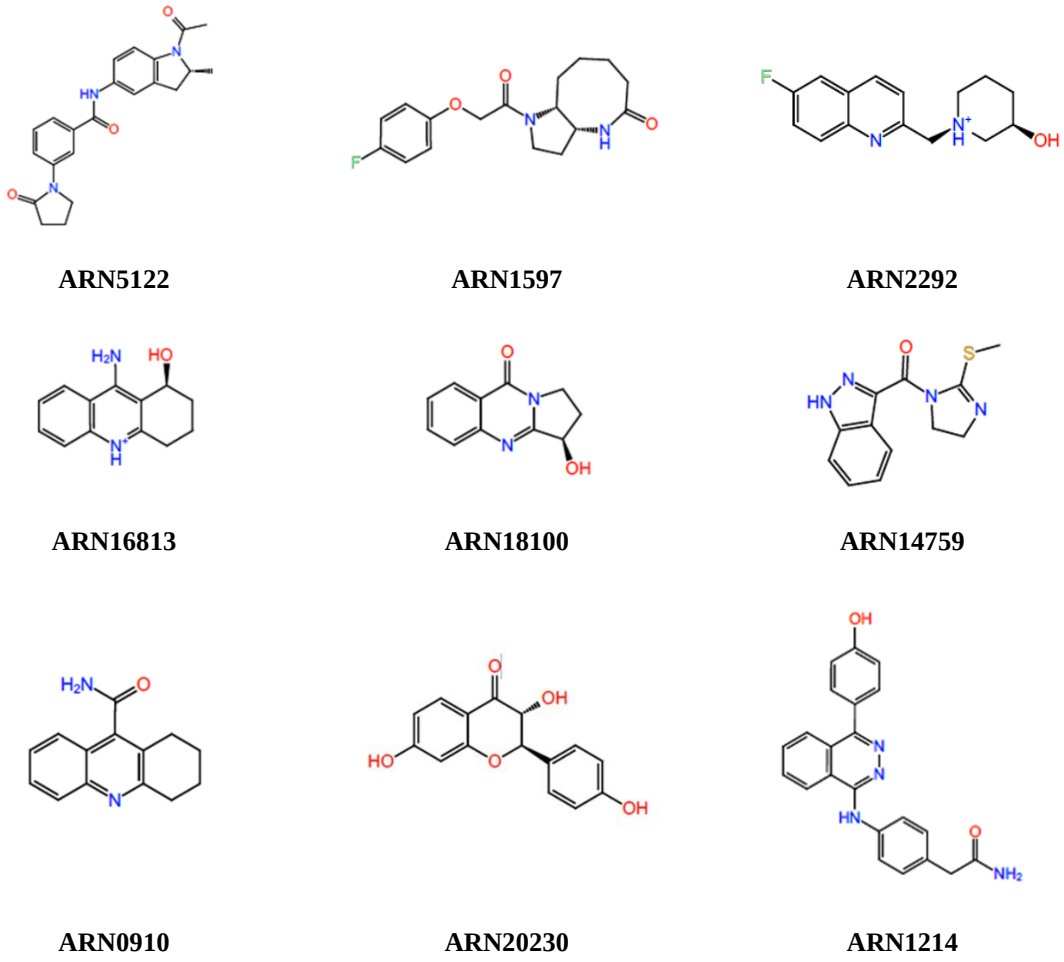


Figure 4. 34. Chemical structures of virtual screening hits

In order to experimentally characterize and further confirm the binding between selected hits and GSK-3β, MST experiments were carried out.

4.6.3. MST on Virtual Screening hits

Compounds arising from virtual screening analysis were analyzed through MST in order to investigate their ability to bind GSK-3β. In this study, GSK-3β was labeled with the red dye NT-647 that labels biomolecules on hexahistidine tails. Among the 9 hits identified with virtual screening, 4 were not available for testing. The solubility of the other 5 compounds was tested with NMR: compound ARN1214 produced aggregates even at low concentrations, and therefore it was excluded

from the analysis. ARN5122, ARN14759, ARN16813 and ARN1597 were tested in presence of saturating concentrations of AMP-PNP with MST, and a preliminary binding check confirmed the binding of all compounds through a good signal to noise ratio. When tested for binding affinity, however, compound ARN14759 did not display a dose-response behavior.

Binding affinity assay with ARN5122 provided a K_D value of 72.8 µM, while ARN16813 provided a K_D of 3.6 µM. The dose-response curves are reported in Figure 4. 35. In binding affinity assays, the change in thermophoresis is expressed as a normalized fluorescence (F_{norm}) which is plotted against molecular concentration to obtain a binding curve and consequently the K_D value of the interaction.

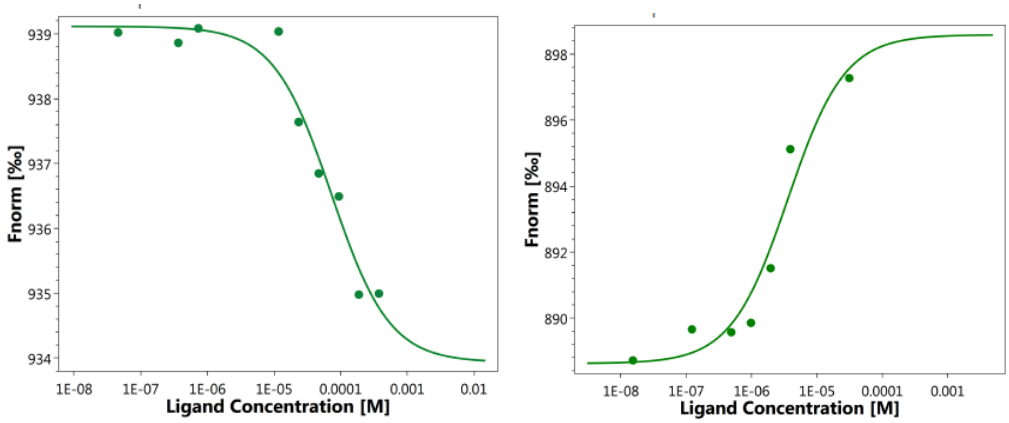


Figure 4. 35. Left: MST analysis of GSK-3β-ARN5122 binding (K_D=72.8µM). Right: MST analysis of GSK-3β-ARN16813 binding (K_D=3.6 µM).

Tested with binding affinity assay, compound ARN1597 displayed a dose-response binding and a K_D value of 28.9 µM. This result further confirmed the binding to GSK-3β previously observed with virtual screening. This compound was also selected by the NMR fragment screening described in the following paragraph. These evidences point out a potential role for ARN1597 as a GSK-3β allosteric modulator. Its action should be also validated by an ATP- and substrate- competition assay and crystallographic studies to verify the presence of actual allosteric inhibition.

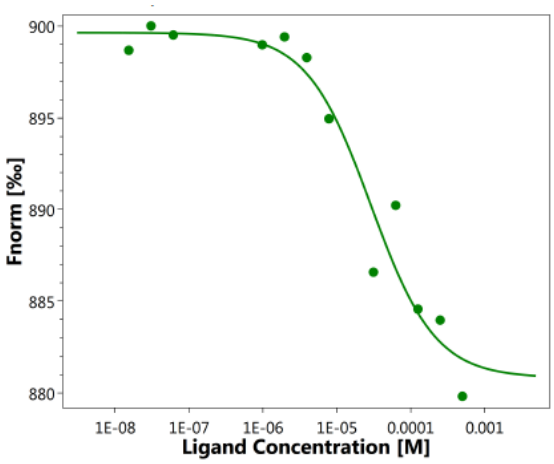


Figure 4. 36. MST analysis of GSK-3β-ARN1597 binding. The graph displays a titration curve of NT647- GSK-3β with increasing concentrations of ARN1597.

4.6.4. Fragment-based NMR screening

Fragment Based Approach (FBA) is a screening method derived in the '80s developed to overcome High Throughput Screening limitations. The advantage of this technique relies on the investigation of the chemical space of a binding site by using small molecules instead of bigger chemical entities; in fact, small molecules can better fit the different binding pockets. The hit rate for screening fragment is, however, significantly higher than for bigger molecules.

Hits coming from FBA usually show weak affinity for the target protein (in the µM-mM range), due to their chemical simplicity and reduced functionality. For this reason, fragments need to be screened at high concentration. Fragment hits are then grown or combined in a more complex molecule to produce a lead with higher affinity. Several techniques can be applied for FBA (such as X-ray crystallography, SPR and fluorescence spectroscopy), but in the present research we focused only on NMR spectroscopy-based screenings.

In order to discover novel allosteric GSK-3β modulators, we screened a homemade local Environment of Fluorine (LEF) library of IIT constituted by 350 fluorinated compounds ranging from 100 to 350 Da and a solubility buffer higher than 400 µM.

Fragments were screened at a concentration of 40 µM in mixtures of 20 compounds each, in the presence and absence of 1 µM GSK-3β pre-incubated for 10 minutes with 1 mM AMP-PNP. The incubation with saturating concentration of AMP-PNP was performed in order to block the ATP binding site, and select only fragments that bind to a different part of the molecule. The knowledge of the chemical shifts of the fragments allowed to easily identify the binders.

Out of the 350 fragments tested, 17 binders were identified. An example of the ^{19}F-NMR screening performed is depicted in Figure 4. 37, in which the ^{19}F-NMR chemical shifts in presence (red) and absence (black) of GSK-3β are reported. Most of the signals do not show any significant shift in presence of the protein, therefore they are not considered GSK-3β binders. Instead, compounds ARN1594 and ARN1597 showed a broadening of their ^{19}F-NMR signals in presence of GSK-3β, indicating a binding to the protein. Interestingly, these compounds were also selected as potential allosteric binders by virtual screening analysis and ARN1497 binding was confirmed through MST experiments as described in the previous paragraph.

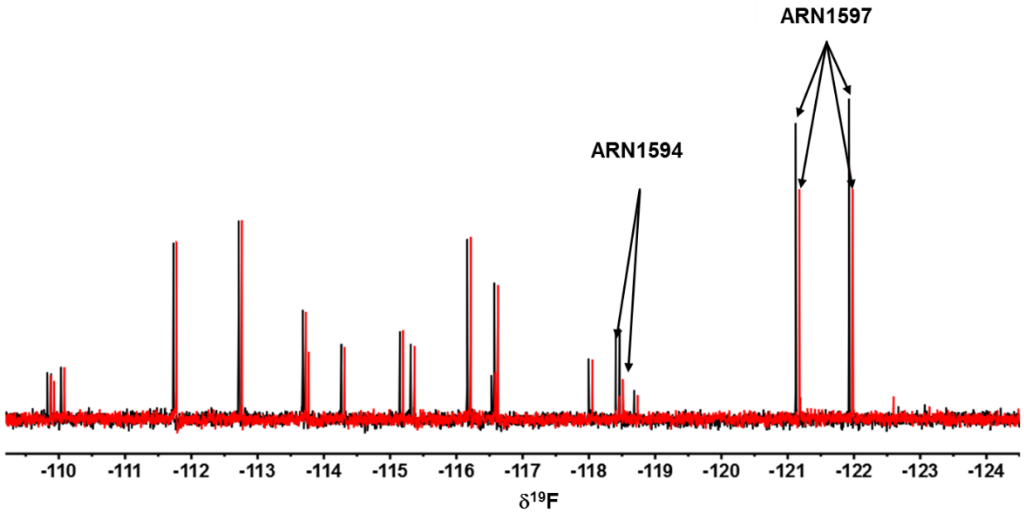


Figure 4. 37. ^{19}F-NMR screening of compounds mixtures in presence (red) and in absence (black) of GSK-3β

The binding was further confirmed by testing the 17 identified hits as single compounds in the same experimental conditions in the presence of a non-binder as internal negative control. Among

these, 14 fragments were confirmed as binders also in presence of 1mM AMP-PNP, even if the effects were mostly weak. Figure 4. 38 shows the ^{19}F-NMR signal of compound ARN1594 tested alone.

All fragment hits will be tested with MST in order to calculate their binding affinity to GSK-3β.

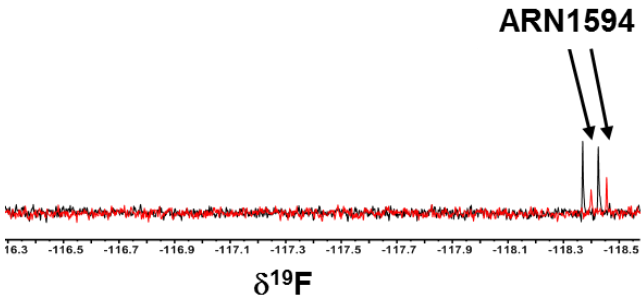
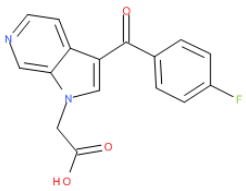
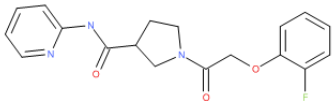


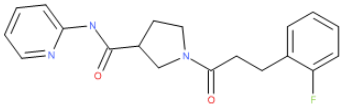
Figure 4. 38. 19F-NMR test of ARN1594 in presence (red) and in absence (black) of GSK-3β



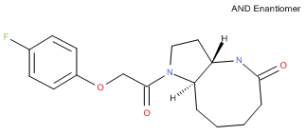
ARN1494



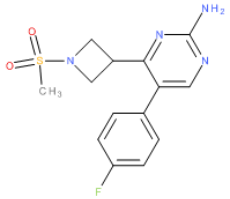
ARN1484



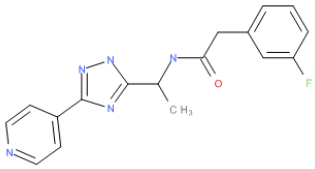
ARN1594



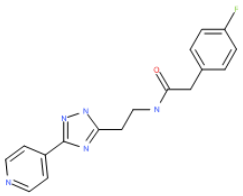
ARN1597



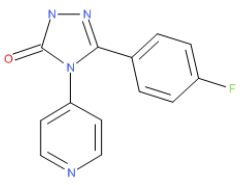
ARN1665



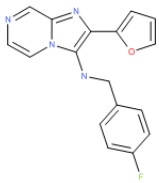
ARN1744



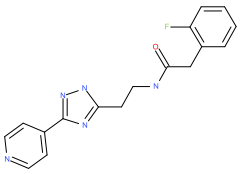
ARN1741



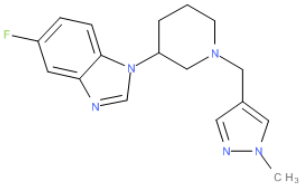
ARN1765



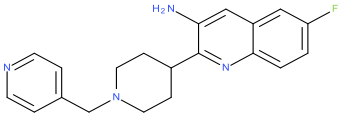
ARN1766



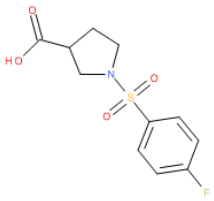
ARN1750



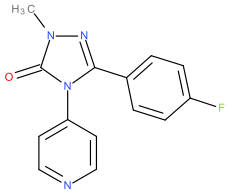
ARN1887



ARN2286



ARN2352



ARN2383

Figure 4. 39. Chemical structures of ^{19}F-NMR screening hits

5. Conclusions and perspectives

GSK-3β (glycogen synthase kinase-3β) is an attractive target for medicinal chemists to design new inhibitors for pathologies that have currently many limitations in therapeutic treatment (such as cancer and neurodegenerative diseases). One of the main challenges in kinase drug discovery is the design of small-molecules inhibitors with selectivity profiles, given the high degree of conservation of kinases ATP-binding sites. The design of allosteric inhibitors and the modulation of drug-target residence time are two different approaches that can be pursued to achieve selectivity. In the present research, we focused on both approaches that were applied in two different drug discovery projects.

The first step was the optimization of a protocol for the expression and purification of full-length GSK-3β. Initially, GSK-3β was expressed in mammalian cells HEK293T, exploiting the novel affinity tag technology relying on the calcium-dependent interaction between the EF1 and EF2 domains of human calbindin D9k. This expression method, however, provided an inadequate protein yield for subsequent experimental purposes. Therefore, we decided to express GSK-3β with his_6-tag employing the baculovirus system in insect cells. The optimal expression conditions were identified in infection for 48h of H5 cells with a multiplicity of infection of 33.3 µL virus/10^6 cells. The protein was purified by two chromatographic steps: an affinity chromatography (NiNTA resin) and a cation-exchange chromatography (HiTrap SP column). Three elution peaks were detected and further characterized: LC-MS/MS analysis highlighted a different phosphorylation pattern, as pool 1 was phosphorylated on Tyr216, while pool 2 and pool 3 were not phosphorylated. These results were in agreement with fluorescent gel shift electrophoresis and 3-FABS NMR activity assays that showed a significantly higher catalytic activity of pool 1, confirming that phosphorylation on Tyr216 enhances GSK-3β activity. The post-translational modification increases also the protein thermal stability, as shown by thermal shift assays. Given these observations, the phosphorylated pool of GSK-3β was employed for all subsequent experiments. The characterization of purified GSK-3β activity was also carried out with NMR by measuring the K_M of ATP (55 µM) and the K_i of staurosporine (7.4 nM). Both values were remarkably coherent with the values reported in the literature.

After optimizing the protocol for GSK-3β expression and characterizing the protein's activity and stability, purified GSK-3β was employed for SPR experiments. The project, performed in collaboration with our computational group, aimed at testing the predictivity of a novel computational algorithm based on adiabatic-bias molecular dynamics (ABMD) simulations. The dissociation process of a series of congeneric GSK-3β ATP-competitive inhibitors showing different experimental potencies were simulated and analyzed with SPR assays.

Suitable immobilization conditions were found by employing a capture-coupling approach, which provided a stable baseline for binding experiments. The identification of an optimal regeneration solution turned out to be a very challenging step, and required many different trials. A solution of 15% DMSO in H_2O allowed the complete dissociation of the bound compounds while preserving the activity of the immobilized protein. Using optimized immobilization and regeneration conditions, we were able to characterize the binding kinetics of the selected ATP-competitive pyrazine inhibitors, which were synthetized by our chemistry group. In order to provide reliable starting configurations for computational simulations, X-ray co-crystal structures of compounds C22, C44 and C50 bound to GSK-3β were generated. Binding poses confirmed the localization in the ATP binding pocket and highlighted a very similar orientation of the inhibitors inside the pocket. Data were acquired at Elettra Synchrotron in Trieste, and structures of C22, C44 and C50 were obtained with good resolution (2.5Å, 2.3Å and 3.2Å, respectively). Coordinates were deposited in the PDB with IDs 6HK3, 6HK4, 6HK7. Unbinding-time-based ranking correlations resulted from elABMD simulations were coherent with experimental data obtained with SPR assays, confirming the validity of the novel algorithm in the qualitative description of the unbinding mechanism in drug discovery programs.

Two benzothiazinones compounds that displayed allosteric inhibition of GSK-3β, 20g and 20f, were selected from the literature (46) and synthetized by our chemistry group. Thermal shift assay (TSA) displayed a very subtle increase of GSK-3β melting temperature upon incubation with these inhibitors, suggesting a very low affinity of binding. A binding check experiment on 20g performed after incubation with AMP-PNP highlighted a signal to noise ratio compatible with the binding to GSK-3β, but displayed the presence of compound aggregation. This issue was also observed when

testing 20f. Unfortunately, the attempts of characterizing the binding affinity of these compounds with MST failed because of the low solubility that these molecules showed in aqueous buffer.

Crystallization trials were also pursued by adding the compounds directly on the crystal of the apo protein, in the attempt to overcome the low solubility issue. Different forms of GSK-3β (phosphorylated, unphosphorylated and unphosphorylated without the his-tag) were investigated, and we observed that the best conditions were obtained with the unphosphorylated form without the his-tag. However, despite the good quality of the diffraction patterns, electron densities corresponding to the inhibitors were not detected on GSK-3β surface.

In parallel, we focused on the discovery of novel allosteric inhibitors through biophysical and computational techniques. An analysis of the pockets on GSK-3β surface performed with the algorithm Pocketron revealed the presence of multiple pockets and a coherence with a previous study reported in the literature by Palomo (36). In particular, cavity 5 (corresponding to Pocketron pockets 1 and 10) was identified as palinurin binding site and was selected for virtual screening analysis.

This analysis selected 9 compounds that were further tested with Microscale thermophoresis (MST) after incubation of GSK-3β with AMP-PNP. Four molecules were tested, as the other five were either not available or displayed aggregation at low concentrations. Among these four compounds, ARN14759 did not display a dose-response behavior. ARN5122 and ARN16813 confirmed the binding to GSK-3β and displayed a K_D value of 78.2 μM and 3.6 μM, respectively. ARN1597 displayed a K_D value of 28.9 μM, and it was particularly interesting as it was selected as a hit also by the NMR fragment based drug discovery (FBDD) screening that was carried out in parallel.

In this study, a homemade library of fluorinated compounds was screened by NMR for the identification of molecules able to bind GSK-3β in a different site from the orthosteric site. Out of 350 fragments screened, 14 were confirmed through NMR experiments as GSK-3β binders in presence of 1 mM AMP-PNP. These compounds will be further characterized by MST experiments.

In summary, the present study has allowed the validation of a promising novel computational algorithm that could become an efficient and innovative tool for calculating residence times in drug discovery programs, offering a qualitative description of drug-target unbinding mechanism. Moreover, we have set up a fast and efficient protocol for full-length GSK-3β expression in insect

cells that allows the separation of different phosphorylation states. We also generated three novel X-ray crystal structures of pyrazine inhibitors bound to the orthosteric site of GSK-3β. Through biophysical and computational techniques, we have explored GSK-3β surface and screened our in-house libraries looking for novel allosteric modulators. In particular, we have identified three promising fragments (ARN1597, ARN5122 and ARN16813) that display affinity for GSK-3β in the micromolar range.

Despite being a lower affinity compared to most ATP-competitive inhibitors, it might be worth taking these compounds into consideration for further developments, as a strong inhibition of GSK-3β has raised many concerns regarding its potential toxicity in therapeutic treatment. In order to confirm the presence of an allosteric regulation, further experiments will be performed. NMR activity assays in presence of increasing concentrations of the molecules will assess whether the catalytic activity of GSK-3β is affected by the binding of the fragments, and ATP- and substrate-competition assays will clarify if the protein inhibition is induced by an allosteric mechanism. Taking advantage of the high solubility of chemical fragments, SPR assays will be carried out to characterize the kinetic of binding and protein-ligand binding affinity.

Considering that to date no crystal structure of an allosteric inhibitor bound to GSK-3β has been described yet, it is worthwhile putting efforts into the investigation of the structural basis of this type of modulation via X-ray crystallography experiments. The identification of the determinants involved in the allosteric inhibition mechanism will provide useful insights for future rational drug design and discovery of small selective GSK-3β inhibitors, contributing to the development of treatments for many severe pathologies.

References

1. Cohen P, Frame S. The renaissance of GSK3. Nat Rev Mol Cell Biol. 2001;2(10):769–776.
2. Pandey MK, DeGrado TR. Glycogen Synthase Kinase-3 (GSK-3)-Targeted Therapy and Imaging. Theranostics. 2016;6(4):571–93.
3. Palomo V, Perez DI, Roca C, Anderson C, Rodriguez-Muela N, Perez C, et al. Subtly Modulating Glycogen Synthase Kinase 3: Allosteric Inhibitor Development and Their Potential for the Treatment of Chronic Diseases. J Med Chem. 2017 Jun 22;60(12):4983–5001.
4. Wagner FF, Bishop JA, Gale JP, Shi X, Walk M, Ketterman J, et al. Inhibitors of Glycogen Synthase Kinase 3 with Exquisite Kinome-Wide Selectivity and Their Functional Effects. ACS Chem Biol [Internet]. 2016 May 13 [cited 2016 Jun 16]; Available from: http://pubs.acs.org/doi/abs/10.1021/acschembio.6b00306
5. Domoto T, Pyko IV, Furuta T, Miyashita K, Uehara M, Shimasaki T, et al. Glycogen synthase kinase-3β is a pivotal mediator of cancer invasion and resistance to therapy. Cancer Sci. 2016 Oct;107(10):1363–72.
6. Walz A, Ugolkov A, Chandra S, Kozikowski A, Carneiro BA, O'Halloran TV, et al. Molecular Pathways: Revisiting Glycogen Synthase Kinase-3β as a Target for the Treatment of Cancer. Clin Cancer Res. 2017 Apr 15;23(8):1891–7.
7. Maqbool M, Mobashir M, Hoda N. Pivotal role of glycogen synthase kinase-3: A therapeutic target for Alzheimer's disease. Eur J Med Chem. 2016 Jan;107:63–81.
8. Mancinelli R, Carpino G, Petrungaro S, Mammola CL, Tomaipitinca L, Filippini A, et al. Multifaceted Roles of GSK-3 in Cancer and Autophagy-Related Diseases. Oxid Med Cell Longev. 2017;2017:1–14.
9. Li D-W, Liu Z-Q, Wei-Chen, Min-Yao, Li G-R. Association of glycogen synthase kinase-3β with Parkinson's disease (Review). Mol Med Rep. 2014 Jun;9(6):2043–50.
10. Sutherland C. What Are the bona fide GSK3 Substrates? Int J Alzheimers Dis. 2011;2011:1–23.
11. Beurel E, Grieco SF, Jope RS. Glycogen synthase kinase-3 (GSK3): Regulation, actions, and diseases. Pharmacol Ther. 2015 Apr;148:114–31.
12. ter Haar E, Coll JT, Austen DA, Hsiao H-M, Swenson L, Jain J. Structure of GSK3β reveals a primed phosphorylation mechanism. Nat Struct Mol Biol. 2001 Jul;8(7):593–6.
13. Stamos JL, Chu ML-H, Enos MD, Shah N, Weis WI. Structural basis of GSK-3 inhibition by N-terminal phosphorylation and by the Wnt receptor LRP6. Elife. 2014;3:e01998.
14. Medina M, Wandosell F. Deconstructing GSK-3: The Fine Regulation of Its Activity. Int J Alzheimers Dis. 2011;2011:1–12.

15. Frame S, Cohen P. GSK3 takes centre stage more than 20 years after its discovery. Biochem J. 2001;359(1):1–16.
16. Dajani R, Fraser E, Roe SM, Young N, Good V, Dale TC, et al. Crystal Structure of Glycogen Synthase Kinase 3β: Structural Basis for Phosphate-Primed Substrate Specificity and Autoinhibition. Cell. 2001 Jun 15;105(6):721–32.
17. Buch I, Fishelovitch D, London N, Raveh B, Wolfson HJ, Nussinov R. Allosteric Regulation of Glycogen Synthase Kinase 3β: A Theoretical Study. Biochemistry. 2010 Dec 28;49(51):10890–901.
18. Bax B, Carter PS, Lewis C, Guy AR, Bridges A, Tanner R, et al. The structure of phosphorylated GSK-3β complexed with a peptide, FRATtide, that inhibits β-catenin phosphorylation. Structure. 2001;9(12):1143–1152.
19. Hernández F, Gómez de Barreda E, Fuster-Matanzo A, Lucas JJ, Avila J. GSK3: A possible link between beta amyloid peptide and tau protein. Exp Neurol. 2010 Jun;223(2):322–5.
20. Hooper C, Killick R, Lovestone S. The GSK3 hypothesis of Alzheimer's disease: GSK3 and Alzheimer's disease. J Neurochem. 2008 Mar;104(6):1433–9.
21. King MK, Pardo M, Cheng Y, Downey K, Jope RS, Beurel E. Glycogen synthase kinase-3 inhibitors: Rescuers of cognitive impairments. Pharmacol Ther. 2014 Jan;141(1):1–12.
22. Martinez A. Preclinical efficacy on GSK-3 inhibitors: Towards a future generation of powerful drugs. Med Res Rev. 2008 Sep;28(5):773–96.
23. Palomo V, I Perez D, Gil C, Martínez A. The potential role of glycogen synthase kinase 3 inhibitors as amyotrophic lateral sclerosis pharmacological therapy. Curr Med Chem. 2011;18(20):3028–3034.
24. Koh S-H, Kim Y, Kim HY, Hwang S, Lee CH, Kim SH. Inhibition of glycogen synthase kinase-3 suppresses the onset of symptoms and disease progression of G93A-SOD1 mouse model of ALS. Exp Neurol. 2007 Jun;205(2):336–46.
25. McCubrey JA, Steelman LS, Bertrand FE, Davis NM, Sokolosky M, Abrams SL, et al. GSK-3 as potential target for therapeutic intervention in cancer. Oncotarget [Internet]. 2014 May 30 [cited 2019 May 13];5(10). Available from: http://www.oncotarget.com/fulltext/2037
26. Tejeda-Muñoz N, Robles-Flores M. Glycogen synthase kinase 3 in Wnt signaling pathway and cancer: GSK-3 in Wnt Signaling Pathway and Cancer. IUBMB Life. 2015 Dec;67(12):914–22.
27. Bilim V, Ougolkov A, Yuuki K, Naito S, Kawazoe H, Muto A, et al. Glycogen synthase kinase-3: a new therapeutic target in renal cell carcinoma. Br J Cancer. 2009 Dec;101(12):2005–14.
28. Naito S, Bilim V, Yuuki K, Ugolkov A, Motoyama T, Nagaoka A, et al. Glycogen Synthase Kinase-3β: A Prognostic Marker and a Potential Therapeutic Target in Human Bladder Cancer. Clin Cancer Res. 2010 Nov 1;16(21):5124–32.

29. Eldar-Finkelman H, Martinez A. GSK-3 Inhibitors: Preclinical and Clinical Focus on CNS. Front Mol Neurosci [Internet]. 2011 [cited 2016 Jan 12];4. Available from: http://journal.frontiersin.org/article/10.3389/fnmol.2011.00032/abstract
30. Kramer T, Schmidt B, Lo Monte F. Small-Molecule Inhibitors of GSK-3: Structural Insights and Their Application to Alzheimer's Disease Models. Int J Alzheimers Dis. 2012;2012:1–32.
31. Ortega MA, Montoya ME, Zarranz B, Jaso A, Aldana I, Leclerc S, et al. Pyrazolo[3,4- b]quinoxalines. A new class of cyclin-Dependent kinases inhibitors. Bioorg Med Chem. 2002 Jul;10(7):2177–84.
32. Chioua M, Samadi A, Soriano E, Lozach O, Meijer L, Marco-Contelles J. Synthesis and biological evaluation of 3,6-diamino-1H-pyrazolo[3,4-b]pyridine derivatives as protein kinase inhibitors. Bioorg Med Chem Lett. 2009 Aug;19(16):4566–9.
33. Logé C, Testard A, Thiéry V, Lozach O, Blairvacq M, Robert J-M, et al. Novel 9-oxo-thiazolo[5,4-f]quinazoline-2-carbonitrile derivatives as dual cyclin-dependent kinase 1 (CDK1)/glycogen synthase kinase-3 (GSK-3) inhibitors: Synthesis, biological evaluation and molecular modeling studies. Eur J Med Chem. 2008 Jul;43(7):1469–77.
34. Bhat R, Xue Y, Berg S, Hellberg S, Ormo M, Nilsson Y, et al. Structural Insights and Biological Effects of Glycogen Synthase Kinase 3-specific Inhibitor AR-A014418. J Biol Chem. 2003 Nov 14;278(46):45937–45.
35. Berg S, Bergh M, Hellberg S, Högdin K, Lo-Alfredsson Y, Söderman P, et al. Discovery of Novel Potent and Highly Selective Glycogen Synthase Kinase-3β (GSK3β) Inhibitors for Alzheimer's Disease: Design, Synthesis, and Characterization of Pyrazines. J Med Chem. 2012 Nov 8;55(21):9107–19.
36. Palomo V, Soteras I, Perez DI, Perez C, Gil C, Campillo NE, et al. Exploring the Binding Sites of Glycogen Synthase Kinase 3. Identification and Characterization of Allosteric Modulation Cavities. J Med Chem. 2011 Dec 22;54(24):8461–70.
37. Wu P, Clausen MH, Nielsen TE. Allosteric small-molecule kinase inhibitors. Pharmacol Ther. 2015 Dec;156:59–68.
38. Martinez A, Alonso M, Castro A, Pérez C, Moreno FJ. First Non-ATP Competitive Glycogen Synthase Kinase 3 β (GSK-3β) Inhibitors: Thiadiazolidinones (TDZD) as Potential Drugs for the Treatment of Alzheimer's Disease. J Med Chem. 2002 Mar;45(6):1292–9.
39. Palomo V, Perez DI, Perez C, Morales-Garcia JA, Soteras I, Alonso-Gil S, et al. 5-Imino-1,2,4-Thiadiazoles: First Small Molecules As Substrate Competitive Inhibitors of Glycogen Synthase Kinase 3. J Med Chem. 2012 Feb 23;55(4):1645–61.
40. Perez DI, Conde S, Pérez C, Gil C, Simon D, Wandosell F, et al. Thienylhalomethylketones: Irreversible glycogen synthase kinase 3 inhibitors as useful pharmacological tools. Bioorg Med Chem. 2009 Oct;17(19):6914–25.

41. Hamann M, Alonso D, Martín-Aparicio E, Fuertes A, Pérez-Puerto MJ, Castro A, et al. Glycogen Synthase Kinase-3 (GSK-3) Inhibitory Activity and Structure–Activity Relationship (SAR) Studies of the Manzamine Alkaloids. Potential for Alzheimer's Disease. J Nat Prod. 2007 Sep;70(9):1397–405.
42. Bidon-Chanal A, Fuertes A, Alonso D, Pérez DI, Martínez A, Luque FJ, et al. Evidence for a new binding mode to GSK-3: Allosteric regulation by the marine compound palinurin. Eur J Med Chem. 2013 Feb;60:479–89.
43. Peng J, Kudrimoti S, Prasanna S, Odde S, Doerksen RJ, Pennaka HK, et al. Structure−Activity Relationship and Mechanism of Action Studies of Manzamine Analogues for the Control of Neuroinflammation and Cerebral Infections. J Med Chem. 2010 Jan 14;53(1):61–76.
44. Zhang P, Hu H-R, Huang Z-H, Lei J-Y, Chu Y, Ye D-Y. Identification of novel scaffold of benzothiazepinones as non-ATP competitive glycogen synthase kinase-3β inhibitors through virtual screening. Bioorg Med Chem Lett. 2012 Dec;22(23):7232–6.
45. Zhang P, Li S, Gao Y, Lu W, Huang K, Ye D, et al. Novel benzothiazinones (BTOs) as allosteric modulator or substrate competitive inhibitor of glycogen synthase kinase 3β (GSK-3β) with cellular activity of promoting glucose uptake. Bioorg Med Chem Lett. 2014 Dec;24(24):5639–43.
46. Gao Y, Ye D-Y, Zhou W-C, Chu Y. The discovery of novel benzothiazinones as highly selective non-ATP competitive glycogen synthase kinase 3β inhibitors for the treatment of ovarian cancer. Eur J Med Chem. 2017 Jul;135:370–81.
47. Ferruz N, De Fabritiis G. Binding Kinetics in Drug Discovery. Mol Inform. 2016 Jul;35(6–7):216–26.
48. Pan AC, Borhani DW, Dror RO, Shaw DE. Molecular determinants of drug–receptor binding kinetics. Drug Discov Today. 2013 Jul;18(13–14):667–73.
49. Copeland RA. Conformational adaptation in drug–target interactions and residence time. Future Med Chem. 2011 Sep;3(12):1491–501.
50. Willemsen-Seegers N, Uitdehaag JCM, Prinsen MBW, de Vetter JRF, de Man J, Sawa M, et al. Compound Selectivity and Target Residence Time of Kinase Inhibitors Studied with Surface Plasmon Resonance. J Mol Biol. 2017 Feb;429(4):574–86.
51. Guo D, Hillger JM, IJzerman AP, Heitman LH. Drug-Target Residence Time-A Case for G Protein-Coupled Receptors: DRUG-TARGET RESIDENCE TIME. Med Res Rev. 2014 Jul;34(4):856–92.
52. Copeland RA. The dynamics of drug-target interactions: drug-target residence time and its impact on efficacy and safety. Expert Opin Drug Discov. 2010 Apr;5(4):305–10.
53. Meyer-Almes F-J. Kinetic binding assays for the analysis of protein–ligand interactions. Drug Discov Today Technol. 2015 Oct;17:1–8.

54. Markgren P-O, Schaal W, Hämäläinen M, Karlén A, Hallberg A, Samuelsson B, et al. Relationships between Structure and Interaction Kinetics for HIV-1 Protease Inhibitors. J Med Chem. 2002 Dec;45(25):5430–9.
55. Magotti P, Ricklin D, Qu H, Wu Y-Q, Kaznessis YN, Lambris JD. Structure-kinetic relationship analysis of the therapeutic complement inhibitor compstatin. J Mol Recognit. 2009 Nov;22(6):495–505.
56. Cusack KP, Wang Y, Hoemann MZ, Marjanovic J, Heym RG, Vasudevan A. Design strategies to address kinetics of drug binding and residence time. Bioorg Med Chem Lett. 2015 May;25(10):2019–27.
57. Dickson A, Tiwary P, Vashisth H. Kinetics of Ligand Binding Through Advanced Computational Approaches: A Review. Curr Top Med Chem [Internet]. 2017 Aug 8 [cited 2019 Apr 15];17(23). Available from: http://www.eurekaselect.com/151618/article
58. Tummino PJ, Copeland RA. Residence Time of Receptor−Ligand Complexes and Its Effect on Biological Function. Biochemistry. 2008 May;47(20):5481–92.
59. Copeland RA, Pompliano DL, Meek TD. Drug–target residence time and its implications for lead optimization. Nat Rev Drug Discov. 2006;5(9):730–739.
60. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-POR). :6.
61. Scholz J, Suppmann S. A new single-step protocol for rapid baculovirus-driven protein production in insect cells. BMC Biotechnol [Internet]. 2017 Dec [cited 2019 Jun 13];17(1). Available from: https://bmcbiotechnol.biomedcentral.com/articles/10.1186/s12896-017-0400-3
62. Choi H, Choi B, Seo J, Lee K, Gye M, Kim Y-P. Rapid Detection of Glycogen Synthase Kinase-3 Activity in Mouse Sperm Using Fluorescent Gel Shift Electrophoresis. Sensors. 2016 Apr 16;16(4):551.
63. Day YSN, Baird CL, Rich RL, Myszka DG. Direct comparison of binding equilibrium, thermodynamic, and rate constants determined by surface- and solution-based biophysical methods. Protein Sci. 2002 May;11(5):1017–25.
64. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010 Feb 1;66(2):125–32.
65. Evans PR. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr. 2011 Apr 1;67(4):282–92.
66. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. 2011 Apr 1;67(4):235–42.
67. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser: crystallographic software. J Appl Crystallogr. 2007 Aug 1;40(4):658–74.

68. Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, et al. Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr. 2012 Apr 1;68(4):352–67.
69. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr. 2010 Apr 1;66(4):486–501.
70. Vagin AA, Steiner RA, Lebedev AA, Potterton L, McNicholas S, Long F, et al. REFMAC5 dictionary: organization of prior chemical knowledge and guidelines for its use. Acta Crystallogr D Biol Crystallogr. 2004 Dec 1;60(12):2184–95.
71. Milletti F, Storchi L, Sforna G, Cruciani G. New and Original pKa Prediction Method Using Grid Molecular Interaction Fields. J Chem Inf Model. 2007 Nov;47(6):2172–81.
72. Decherchi S, Bottegoni G, Spitaleri A, Rocchia W, Cavalli A. BiKi Life Sciences: A New Suite for Molecular Dynamics and Related Methods in Drug Discovery. J Chem Inf Model. 2018 Feb 26;58(2):219–24.
73. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013 Apr 1;29(7):845–54.
74. La Sala G, Decherchi S, De Vivo M, Rocchia W. Allosteric Communication Networks in Proteins Revealed through Pocket Crosstalk Analysis. ACS Cent Sci [Internet]. 2017 Aug 10 [cited 2017 Sep 19]; Available from: http://pubs.acs.org/doi/abs/10.1021/acscentsci.7b00211
75. Madhavi Sastry G, Adzhigirey M, Day T, Annabhimoju R, Sherman W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. J Comput Aided Mol Des. 2013 Mar;27(3):221–34.
76. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. J Med Chem. 2004 Mar;47(7):1739–49.
77. Halgren TA. Identifying and Characterizing Binding Sites and Assessing Druggability. J Chem Inf Model. 2009 Feb 23;49(2):377–89.
78. Dunning CJ, McGauran G, Willén K, Gouras GK, O'Connell DJ, Linse S. Direct High Affinity Interaction between Aβ42 and GSK3α Stimulates Hyperphosphorylation of Tau. A New Molecular Link in Alzheimer's Disease? ACS Chem Neurosci. 2016 Feb 17;7(2):161–70.
79. Savitsky P, Bray J, Cooper CDO, Marsden BD, Mahajan P, Burgess-Brown NA, et al. High-throughput production of human proteins for crystallization: The SGC experience. J Struct Biol. 2010 Oct;172(1):3–13.
80. Dalvit C, Ardini E, Flocco M, Fogliatto GP, Mongelli N, Veronesi M. A General NMR Method for Rapid, Efficient, and Reliable Biochemical Screening. J Am Chem Soc. 2003 Nov;125(47):14620–5.

81. Grøftehauge MK, Hajizadeh NR, Swann MJ, Pohl E. Protein–ligand interactions investigated by thermal shift assays (TSA) and dual polarization interferometry (DPI). Acta Crystallogr D Biol Crystallogr. 2015 Jan 1;71(1):36–44.
82. Marchi M, Ballone P. Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems. J Chem Phys. 1999 Feb 22;110(8):3697–702.
83. Rich RL, Errey J, Marshall F, Myszka DG. Biacore analysis with stabilized G-protein-coupled receptors. Anal Biochem. 2011 Feb;409(2):267–72.
84. Drake AW, Klakamp SL. A strategic and systematic approach for the determination of biosensor regeneration conditions. J Immunol Methods. 2011 Aug;371(1–2):165–9.
85. Goode JA, Rushworth JVH, Millner PA. Biosensor Regeneration: A Review of Common Techniques and Outcomes. Langmuir. 2015 Jun 16;31(23):6267–76.
86. Renaud J-P, Chung C, Danielson UH, Egner U, Hennig M, Hubbard RE, et al. Biophysics in drug discovery: impact, challenges and opportunities. Nat Rev Drug Discov. 2016 Oct;15(10):679–98.
87. Cimmperman P, Baranauskienė L, Jachimovičiūtė S, Jachno J, Torresan J, Michailovienė V, et al. A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. Biophys J. 2008 Oct;95(7):3222–31.